DOCUMENT RESUME

ED 024 636 24 SP 001 889

By- Glass, Gene V.; Maguire, Thomas O.

ANALYSIS OF TIME-SERIES QUASI-EXPERIMENTS. FINAL REPORT.

Alberta Univ., Edmonton.; Colorado Univ., Boulder.; Illinois Univ., Urbana. Center for Instructional Research and Curriculum Evaluation (CIRCE).
Spons Agency- Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No- BR-6-8329
Pub Date 24 Mar 68
Grant- OEG-3-7-008329-2065
Note- 122p.
EDRS Price MF-$0.50 HC-$6.20
Descriptors- Computer Programs, Control Groups, Educational Psychology, *Educational Research, Experiments, *Models, Psychology, *Research Design, *Research Methodology, *Statistical Analysis
Identifiers- Time-Series Quasi-Experiments

The objective of this project was to investigate the adequacy of statistical models developed by G. E. P. Box and G. C. Tiao for the analysis of time-series quasi-experiments: (1) The basic model developed by Box and Tiao is applied to actual time-series experiment data from two separate experiments, one in psychology and one in educational psychology. (2) The same model is applied in a relatively complex experimental design to analyze data on traffic fatalities in the state of Connecticut before and after a legislative crackdown on speeding. (3) A generalization by Tiao of the basic model (as yet unpublished) is applied to data on the effects of a revision of Germany's divorce laws in 1900 on the rate of divorce; this model incorporates a "drift" parameter which accommodates series which show either a rise or fall over time. (4) A model which is a particular generalization of the "drift" model is presented; this more general model (developed for this study by Tiao) incorporates a parameter to account for instantaneous change in the direction of the drift of a series associated with the introduction of a treatment. Appendixes A and B contain printouts of computer programs for analyses illustrating (3) and (4). (Author/SG)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BR-6-8329
PA-24

ED0 24636

Final Report

Project No. 6-8329
Grant No. OEG-3-7-008329-2065

Analysis of Time-Series Quasi-Experiments

Gene V Glass
Laboratory of Educational Research
University of Colorado
Boulder, Colorado

Thomas O. Maguire
University of Alberta
Edmonton, Alberta

March 24, 1968

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

SP001889

Table of Contents

Acknowledgments ii

Chapter I. Introduction 1

Chapter II. Analysis of a Class of Time-Series Quasi-Experiments with the Box-Tiao Model 9

Chapter III. Analysis of Data on the Connecticut Speeding Crackdown as a Time-Series Quasi-Experiment 26

Chapter IV. Analysis of Data on the Revision of German Divorce Laws as a Time-Series Quasi-Experiment 62

Chapter V. An Analytic Technique Employing a Change in Drift of a Time-Series Associated with the Introduction of a Treatment 92

Appendix A. Computer Program for Analysis of Time-Series Experiment with Constant Drift 100

Appendix B. Computer Program for Analysis of Time-Series Experiment with Possible Change in Drift 101

Acknowledgments

The research reported herein was partially supported by a grant from the U.S. Office of Education to the University of Illinois. The bulk of the work was carried out while the authors were in residence at the Center for Instructional Research and Curriculum Evaluation of the University of Illinois, the principal investigator as a staff member of that unit and the co-investigator as a graduate student. In June 1967, we both left the University of Illinois for our respective positions. The University of Colorado and the University of Alberta have contributed to our efforts in the final stages of this work; however, these institutions -- and not the University of Illinois -- are credited on the title page of this report only because of explicit directions from the funding agency to list institutional affiliations at the date of submission of the final report.

Three colleagues made substantial contributions to our work. Dr. Donald T. Campbell of the Department of Psychology, Northwestern University, served us as a consultant and supplied data for the analyses which comprise Chapter III. Dr. David E. Wiley of the Department of Education, University of Chicago, contributed valuable insights as a consultant to the project.

The greatest debt is owed to Dr. George C. Tiao of the Department of Statistics, University of Wisconsin. Dr. Tiao responded to our earnest requests for assistance with unselfish concern and with an intellectual competence which is amply displayed in the analytic techniques he derived specifically for Chapters IV and V. In responding with ingenuity and sophistication to a practical need in data analysis, Dr. Tiao has performed in the manner of the _significant_ statisticians of our day and has won our sincere gratitude.

Numerous graduate students contributed to our efforts over the past two years; among them were Thomas Bligh, Donald Bosshart, and Perc D. Peckham. The typing of most of the final report was done by Linda Lou Schmale.

Gene V Glass

Thomas O. Maguire

March 24, 1968
Boulder, Colorado

Chapter I

INTRODUCTION

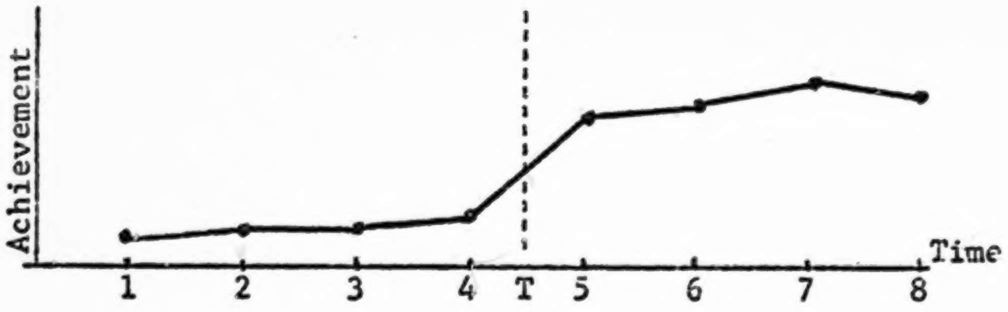
The publication of "Experimental and Quasi-Experimental Designs for Research on Teaching" by D. T. Campbell and J. C. Stanley in the Handbook of Research on Teaching (1963) represents a major advance in educational-research methodology. Already this chapter has become a much-used reference in numerous departments of psychology and sociology as well as in schools of education. The Campbell-Stanley chapter was a presentation of various designs for the assessment or comparison of treatment effects with a discussion of their strengths and weaknesses. "In general, [the multiple-group time-series design] is an excellent quasi-experimental design, perhaps the best of the more feasible designs.... The availability of repeated measurements makes [the design] particularly appropriate to research in schools" (Campbell and Stanley, 1963, p. 227). To date, appropriate inferential statistical analyses of treatment effects in all types of time-series experiments have not been developed. Campbell and Stanley (1963) and Campbell (1963) have lamented the lack of appropriate analytic techniques for important time-series quasi-experimental designs.

Recent developments in mathematical statistics by Box and Jenkins (1962) and Box and Tiao (1965) provide statistical models which may very well fill the need identified by Campbell and Stanley. Broadly conceived, the objective of this project was the investigation of the adequacy of statistical models developed by Box and Tiao for the analysis of time-series quasi-experiments. This investigation involved the following: (1) the investigation of the Box-Tiao models as to their adequacy as descriptions of time-series experimental data, (2) investigation of the possibility of extension of the models of Box and Tiao to the analysis of more general classes of time-series quasi-experiments, (3) the development of computer programs for statistical analysis based on the models, (4) the application of the models to the analysis of actual time-series quasi-experiments.

The Time-Series Experiment

The time-series experiment was identified and discussed at length by Campbell and Stanley (1963) in Gage's _Handbook of Research on Teaching_ and by Campbell (1963). This singularly useful design for experimental research has long been a paradigm for experimentation in the physical sciences.

In the social sciences, and education, a person or group of persons might be observed and measured at regular intervals prior to the introduction of an experimental treatment (T). Several observations on the group follow the introduction of T. An abrupt change in the level of the average score for the group between the observation immediately preceding T and those following it may indicate a cause and effect relationship between T and the variable being measured. Note how in the figure below the introduction of T (a new curriculum perhaps) appears to have increased the "achievement" of the class.



A multiple-group time-series quasi-experiment might involve three treatments (T_1, T_2, T_3), each applied simultaneously to a different group. One would then compare the three gains between times 4 and 5. Another modification of the basic time-series design might be called a dependent-groups time-series design. In such designs, the same group of persons is involved in multiple time-series designs for the evaluation of more than one treatment (T) effect.

Analysis of Time-Series Experiments

A crucial problem, identified but not resolved by Campbell and Stanley, is that of the decision whether the gain in "achievement" between times 4 and 5 should be attributed to random fluctuations in the time-series curve or to an outside influence (presumably T, the new curriculum). This is an inferential statistical problem. Campbell and Stanley (1963) considered several significance-testing procedures for analysis of the time-series design; in the end, each procedure suffered either from implausible assumptions or weaknesses such as disregard for most of the data or lack of power (i.e., failure to produce significant results when any reasonable observer would attribute an effect to T).

A simple correlated _t_-test of the difference between the pre-treatment and post-treatment observations was judged inappropriate by Campbell and Stanley because it would produce significant results for nonstationary time-series in which no abrupt change occurred. A test for the deviation of a value from the regression line derived on all observations preceding it (Mood, 1950, pp. 297-298) was considered.

The dependence of this test on the unrealistic (for time-series analysis) assumption of independent observations makes it inappropriate. By far the greatest amount of research into the problem of analysis of time-series experiments has been accomplished under the direction of Donald T. Campbell at Northwestern University. Much of the work of Campbell and his associates has not yet been published.

Box and Tiao (1965) presented a statistical model for the analysis of the change in level of a nonstationary time-series. The problem they considered was that of making inferences about a possible shift in the level of a time-series associated with the occurrence of an event, which we have called T. This model (an integrated moving average model) is based on statistical assumptions likely to be met in many time-series quasi-experiments. The integrated moving average model is reported to represent quite well a surprisingly large number of time-series in economics and industry, but it is practically unknown to behavioral scientists and educational researchers.

The mere existence of statistical models and analysis procedures does not guarantee their usefulness for educational or psychological research. The economist and chemist have many statistical models that are of no use to behavioral scientists. Response surface methodology (Box, 1954) was designed (primarily for the chemist) to answer questions about optimal combinations of many factors; the method is an invaluable tool for the chemist. However, due to the generally small ratio of treatment effects to experimental error which plagues researchers in the social sciences, they rarely use response surface methodology. One focus of the research reported here was to determine whether the methods of time-series analysis developed by Box and Tiao suffer from the same inability to distinguish treatment effects when they are small compared to error.

The remainder of this report falls into four chapters and two appendices. In Chapter II, the basic model developed by Box and Tiao (1965) is applied to actual time-series experiment data from two separate experiments, one in psychology and the other in educational psychology. In Chapter III, the same model is applied in a relatively complex experimental design to analyze data on traffic fatalities in the state of Connecticut before and after a legislative crackdown on speeding. A generalization by George C. Tiao of the basic model -- as yet unpublished -- is applied to data on the effects of a revision of Germany's divorce laws in 1900 on the rate of divorce, in Chapter IV. This generalized model incorporates a "drift" parameter which accommodates series which show either a rise or a fall over time. In Chapter V a model which is a particular generalization of the "drift" model in Chapter IV is presented. The most general model in Chapter V incorporates a parameter to account for an instantaneous change in the direction of the drift of a series associated with the introduction of a treatment. The general model of Chapter V was also developed for this study by Dr. Tiao. In Appendices A and B appear print outs of computer programs for the analyses illustrated in Chapters IV and V, respectively.

References

Box, G. E. P. The exploration and exploitation of response surfaces: some general considerations and examples. *Biometrics*, 1954, *10*, 16-60.

Box, G. E. P. and Jenkins, G. M. Some statistical aspects of adaptive optimization and control. *Journal of the Royal Statistical Society, B*, 1962, *24*, 297-343.

Box, G. E. P. and Tiao, G. C. A change in level of a non-stationary time series. *Biometrika*, 1965, *52*, 181-192.

Campbell, D. T. From description to experimentation: interpreting trends as quasi-experiments. Chapter 12 in C. W. Harris, *Problems in Measuring Change*. Madison, Wisconsin: University of Wisconsin Press, 1963.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. Chapter 5, (pp. 171-246) in N. L. Gage (Ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.

Mood, A. *Introduction to the Theory of Statistics*. New York: McGraw-Hill, 1950.

Chapter II

ANALYSIS OF A CLASS OF TIME-SERIES QUASI-EXPERIMENTS WITH THE BOX-TIAO MODEL

In a recent article, Box and Tiao (1965) developed a method of evaluating the change in level between two successive points in time of a non-stationary time-series. Observations z_t are taken at equally spaced time intervals and one wishes to make inferences about a possible shift in level of the time-series associated with the occurrence of an event at a particular point in time. This is precisely the situation described by Campbell and Stanley (1963) as a time-series quasi-experimental design. Several observations are taken before and after the administration of a treatment, T, e.g., $O_1\ O_2\ O_3\ T\ O_4\ O_5\ O_6$. If there is an abrupt shift in the level of a time-series between the third and fourth observations, evidence of a treatment effect may exist. Campbell and Stanley recognized the shortcomings in the statistical tests they suggested as possible analytic techniques (Campbell, 1963; Campbell and Stanley, 1963). The model and statistical techniques developed by Box and Tiao (1965) appear to be the most suitable methods now available which might have application to the analysis of time-series quasi-experiments.

The statistical model underlying the Box-Tiao analysis of change in level of a time-series is the integrated moving average model.

$$z_1 = \alpha_1 \quad \text{and} \quad z_t = L + \gamma \sum_{i=1}^{t-1} \alpha_{t-1} + \alpha_t \qquad (3.1a)^*$$

for the $\underline{n}_1$ observations prior to the introduction of T, and

---

*Equation numbers are the same as those in Box and Tiao (1965).

$$z_t = L + \delta + \gamma \sum_{i=1}^{t-1} \alpha_{t-i} + \alpha_i \qquad (3.1b)$$

for the $\underline{n}_2 = \underline{N} - \underline{n}_1$ observations following T, where:

z_t is the value of the variable observed at time $\underline{t}$,

L is a fixed but unknown location parameter,

γ is a parameter descriptive of the degree of interdependence of the observations in the time-series and takes values $0 < \gamma < 2$,

α_t is a random normal deviate with mean 0 and variance σ^2.

δ is the change in level of the time-series caused by T.

Essentially the model implies that the system is subjected to periodic random shocks (α_t) a proportion (γ) of which are absorbed into the level of the series. Data which conform to the model in (3.1a) evidence the following properties (among others):

1. The graph of the time-series follows an erratic, somewhat random path with slight, but no systematic drifts, trends, or cycles.
2. The correlogram (i.e., the graph of the autocorrelations) of the observations, z_t, does not "die out," (i.e., does not tend systematically toward zero as the lag between values correlated increases) nor does it show cycles characteristic of cyclic time-series.
3. For the $\underline{N}$-1 differences between adjacent observations, $z_t - z_{t-1}$, the lag 1 correlation equals $(\gamma - 1)/[1 + (\gamma - 1)^2]$ and all higher lag correlations equal zero.

By setting $y_1 = z_1$, and $y_t = z_t - \gamma \sum_{j=1}^{t-2} (1 - \gamma)^j \, z_{t-1-j}$, the model can be written as $Y = X\,\theta + e$ where X is defined as an N x 2 matrix of weights as follows:

$$\underline{X}^T = \left[\begin{array}{cccc|cccc} 1 & (1-\gamma) & \ldots & (1-\gamma)^{n_1-1} & (1-\gamma)^{n_1} & \ldots & & (1-\gamma)^{N-1} \\ 0 & 0 & \ldots & 0 & 1 & (1-\gamma) & \ldots & (1-\gamma)^{n_2-1} \end{array} \right]$$

θ is a 2 x 1 vector containing as elements L and δ and e is an N x 1 vector of random normal deviates, $e^T = (\alpha_1 \ldots \alpha_N)$, the elements of which have mean 0 and variance σ^2.

When γ is known, simple least squares estimates of L and δ can be found from the familiar solution to the least-squares normal equations:

$$\hat{\theta} = \begin{bmatrix} \hat{L} \\ \hat{\delta} \end{bmatrix} = (X^T X)^{-1} X^T Y \tag{3.4}$$

Box and Tiao showed that the least squares estimate of δ, namely, $\hat{\delta}$, has a t distribution with N - 2 df when divided by an appropriate estimate of its standard error.

When γ is unknown (as will generally be true) a Bayesian analysis using sample information about γ is used in making inferences about δ. The posterior distribution, $h(\gamma|z)$, of γ given a set of N observations and assuming a uniform prior distribution is known to within a constant of proportionality.

$$h(\gamma|z) \propto \frac{\gamma(2-\gamma)}{\sqrt{[1-(1-\gamma)^{2n_1}]\;[1-(1-\gamma)^{2n_2}]}} (s^2)^{-\frac{1}{2}(N-2)} \tag{5.8}$$

where s^2 is the residual variance and is given by

$$s^2 = \frac{1}{N-2} (Y^T Y - \theta^T X^T X \theta), \tag{3.18}$$

for a given value of γ.

Computing Procedures

The present program performs the following operations:

1. The correlograms (autocorrelations for lag 1 through lag $3n_1/4$ or $3n_2/4$) are calculated for z_t and for $z_t - z_{t-1}$ separately for pretreatment and posttreatment data.
2. The posterior distribution, $h(\gamma|z)$, of γ given the data $\underline{z}$ for a uniform prior distribution is claculated and plotted. The value of $h(\gamma/z)$ is found for 200 values of γ from .01 to 2.00 in steps of .01. For each of these 200 values of γ,
3. $\hat{\delta}$ is calculated,
4. The variance error of $\hat{\delta}$ is calculated
5. The t-statistic equal to the ratio of $\hat{\delta}$ to its standard error is calculated and plotted on a graph which is superimposed on the graph of $h(\gamma|z)$.

Formation of the matrix of weights $\underset{\cdot\cdot}{X}$ in (3.3) raises problems with "underflow." As γ approaches 1.0, successive elements of X get very small. Consequently, when $|x_{t1}| < 1.0 \times 10^{-15}$, the effects of the weights can be considered negligible and each subsequent value of x_{t1} is set equal to zero for that γ. The same problem and treatment apply to x_{t2}.

The vector $\hat{\theta}$ is formed by equation (3.4.) and the residual variance s^2 calculated from formula (3.18). The standard error of $\hat{\delta}$ is given by

$$\hat{\sigma}(\hat{\delta}|\gamma) = \sqrt{\frac{s^2\gamma\ (2-\gamma)\ [1-(1-\gamma)^{2\underline{N}}]\ (N-2)}{[1-(1-\gamma_o)^{2n_1}][1-(1-\gamma_o)^{2n_2}](N-4)}}\ ,$$

The t value for testing the significance of the difference of $\hat{\delta}$ from 0 is given by $\underline{t} = \underline{\hat{\delta}}/\hat{\sigma}(\hat{\delta}|\gamma)$.

Special problems of computer accuracy arise in the calculation of the posterior distribution of γ. When s^2 is large, the posterior distribution can be very small, since $s^{-(N-2)}$ is a multiplicative factor in formula (5.8). To prevent any problems associated with this calculation, all factors are transformed to *$\log_{10}$* prior to the calculation of the ordinates of the posterior distribution, which are then transformed by subtracting the largest value thus obtained. Thus, all values of the posterior distribution are divided by the maximum value. With one slight exception, *antilogs* are taken to restore the values to their original form. The exception is that when the *log* of a given value differs from the *log* of the maximum value by more than 35 (i.e., given value / maximum value $< 10^{-35}$), the value of the ordinate corresponding to the given value is set equal to zero.

Each of the values of $\hat{L}$, $\hat{\delta}$, s^2, $\hat{\sigma}_{(\hat{\delta}|\gamma)}$, t, $h(\gamma|z)$, is stored for each value of γ. (The values of the posterior distribution of γ are rescaled by fitting trapezoids so that the curve has unit area.)

Illustrative Analysis

An illustrative analysis will be performed on data from an experiment by Deese and Carpenter which was adapted for presentation in Brown (1961, pp. 118-119). Two groups of rats were given 24 training trials in running a short alley for food. Group A had been fed wet mash for one hour prior to the experiment; group B had not eaten for 22 hours. After 24 trials the conditions were reversed, group A being deprived of food for 22 hours and group B being fed for one hour prior to a final eight trials. Observations were made of the length of time between start of a trial and a running response for each rat. Observations were converted to logarithms of this latency period for each rat which were then averaged and divided into 1 for both groups. The reciprocals of the average log latencies for groups A and B over 32 trials appear in Figure 1. Data for Group B (high drive followed by low drive state) appear as the solid line; the broken line is for Group A.

Figure 1 about here

The significance of the effect of shifting from a low drive state to a high drive state (Group A) is apparent and would not be enhanced by further statistical analysis. However, the significance of the slight downward shift between the 24th and 25th trials for Group B is worthy of further investigation.

Assuming that the fundamental time-series process generating the series is the same for both Groups A and B, the $\underline{n}_1 = 24$ pretreatment observations for both groups should provide a reasonably good, though somewhat unstable, estimate of the correlograms for raw data, z_t, and differences, $z_t - z_{t-1}$. The $\underline{n}_2 = 8$ posttreatment observations are insufficient in number to add any substantial information concerning the fit of the model. Even the small number of pretreatment observations probably represents fewer than a minimal number from which one may draw inferences about the fit of the model with any confidence. The correlograms for Groups A and B for the data in Figure 1 were calculated. To conserve space these correlograms are not reproduced here. These two sets of autocorrelations evidenced neither cycles nor systematic dampening effects characteristic of time series of types other than moving average series. On the basis of inspection of Figure 1 and the correlograms of the original data we can continue to entertain the model of equations (3.1a) and (3.1b).

The adequacy of the model is further investigated by observing the correlogram of the differences between adjacent observations in a series, i.e., the correlogram of $z_{t+1} - z_t$, $t = 1, \ldots, N-1$. The correlograms were calculated for the pretreatment data in Figure 1. The lag 1 through lag 10 autocorrelations of the differences between successive observations appear in Figure 2.

Figure 2 about here

The correlogram for $z_t - z_{t-1}$ where z_t conforms to the model in (3.1a) should show a lag 1 correlation of $(\gamma - 1)\ [1 + (\gamma - 1)^2]$ and lag $\underline{k}$ autocorrelations of zero ($\underline{k} > 1$). For sample data, an approximation to the standard error of an autocorrelation coefficient due to Bartlett (1946) is available. The slanted, straight lines at the top and bottom of Figure 2 mark off a distance of two standard errors of the autocorrelation coefficient of lag $\underline{k}$, $2\sigma_{r_k}$. Note that only one (lag 9 - Group B) of the 18 autocorrelation coefficients in Figure 2 lies in a region of rejection that the population value of an autocorrelation of lag greater than 1 is zero.

As will be seen later, the maximum likelihood estimates of γ are 0 and .25 for Groups A and B, respectively. The lag 1 autocorrelation for group A is almost equal to the expected (on the basis of the model) value of $(\gamma - 1)/[1 + (\gamma - 1)^2] = -.50$. If the model fits exactly, one would expect a lag 1 autocorrelation for Group B of $(.25 - 1)/[1 + (-.75)^2] = -.48$. The obtained value of $-.28$ does not differ significantly from this expected value. Acknowledging the limited power of these statistical tests and the fact that to accept uncritically a model on the basis of so few observations is largely a matter of faith, we proceed with the analysis of the data for Group B assuming that the model of (3.1a) and (3.1b) holds.

In Figure 3 appear the following: (a) the posterior distribution of γ, which indicates the likely values which γ might assume for these data (the maximum likelihood estimate of γ is found by noting that value of γ under the peak of the curve); (b) the t-statistic used in testing the hypothesis that δ, the shift in level of the series associated with the events between trials 24 and 25, is equal to zero; (c) the values (dotted lines) of t needed for significance at the .15, .10 and .05 levels in testing the hypothesis H_0: $\delta = 0$ against the hypothesis H_1: $\delta < 0$. Note that if $\gamma = .20$, H_0 can be rejected at the .05 level in favor of H_1. If γ is .50 or above (which appears relatively unlikely), H_0 cannot be rejected at the .15 level. Our impression is that the data do indicate a statistically significant shift downward after the 24th trial for Group B.

The data in Figure 4 are the number of times in a 50-minute period in school that a four-year-old hyperactive child changed activities over 28 consecutive days. (Allen et al., 1967) For the last seven days, the child was given verbal, social reinforcement for attending to a single activity for more than a minute. The question: Is there a decrease in activity changes associated with introduction of reinforcement of attending behavior between days 21 and 22?

The decrease of approximately 25 activities in 50 minutes from day 21 to day 22 is less than the "natural" decrease of the time-series between days 3 and 4 and days 14 and 15. Can the activity of the time-series over the last seven days be viewed as the regular progression of the time-series over the first 21 days? An inferential statistical analysis will illuminate these data.

The data in Figure 4 show neither upward nor downward trends nor cycles. The lag 1 autocorrelation coefficient for the differences between successive pre-treatment observations is -.5, which corresponds closely to the expected value when $\gamma = 0$ (which is the maximum likelihood estimate of γ for these data). The other autocorrelations for the differences are not significantly different from zero. Thus the data in Figure 4 appear to conform to the model.

As was pointed out earlier, the objective of the analysis of a possible change in level of a time-series of the integrated moving average type is to obtain a least-squares estimate of δ and a distributional statement about the estimate. Provided γ is known, this objective is relatively easily attained. The least-squares estimate of δ, namely $\hat{\delta}$, and an estimate of its standard error were given in Box and Tiao in their 1965 paper. Under suitable assumptions of normality and independence of errors, $(\hat{\delta} - \delta)/\hat{\sigma}(\hat{\delta})$ has a t-distribution with $n_1 + n_2 - 2$ degrees of freedom. A test of the hypothesis that $\delta = 0$ can be carried out with the test statistic $\underline{t} = \hat{\delta}/\hat{\sigma}(\hat{\delta})$.

However, if γ is not known (as is generally true), it is necessary to obtain information from the sample of N observations about probable values for γ. Either of two strategies might be followed: 1) find the maximum likelihood estimate of γ and estimate $\hat{\delta}$ and $\hat{\sigma}(\hat{\delta})$ using only that maximum likelihood estimate; 2) plot both the likelihood distribution of γ and the value of $\underline{t} = \hat{\delta}/\hat{\sigma}(\hat{\delta})$ against the value of γ as it ranges between 0 and 2 and see if the t-statistic is clearly significant or non-significant over the range of probable values for γ. [The likelihood distribution of γ given the N sample observations is given in Box and Tiao

(1965).] Both strategies will be employed at different points in the analyses to follow.

Of course, γ is unknown. The likelihood distribution of γ given the data is plotted as $h(\gamma|z)$ in Figure 5. The chances are practically nil that γ is above .10; the maximum likelihood estimate of γ is zero.

Over the range of likely values of γ, 0 to .10, the value of $\underline{t}$ for testing the hypothesis that $\delta = 0$ is clearly statistically significant ($\underline{t}$ is never greater than -4.4). Without much question, then, the introduction of the treatment at day 21 worked an effect upon the rate of activity change.

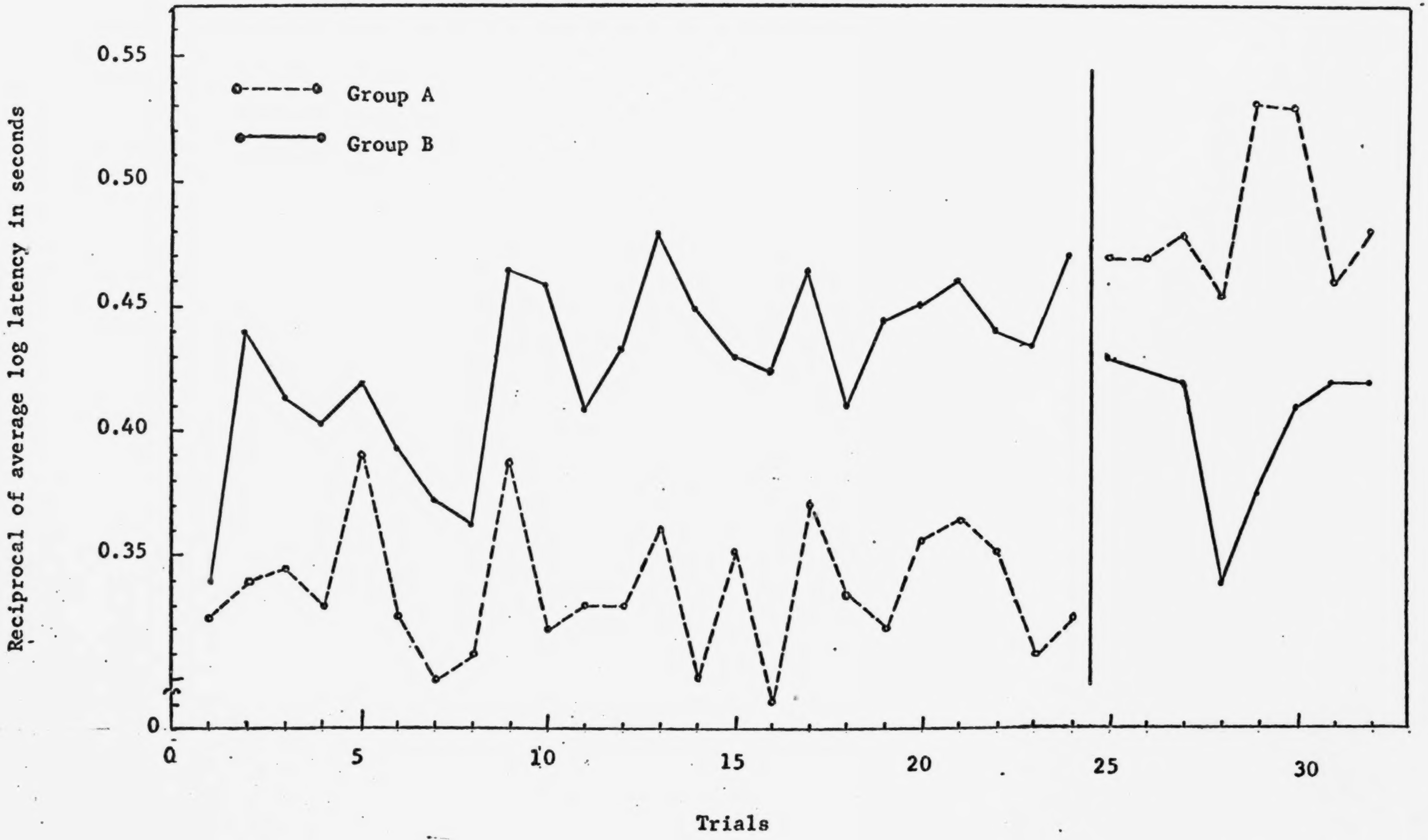


Figure 1. Starting speeds for two groups of rats before and after a reversal in deprivation schedule.

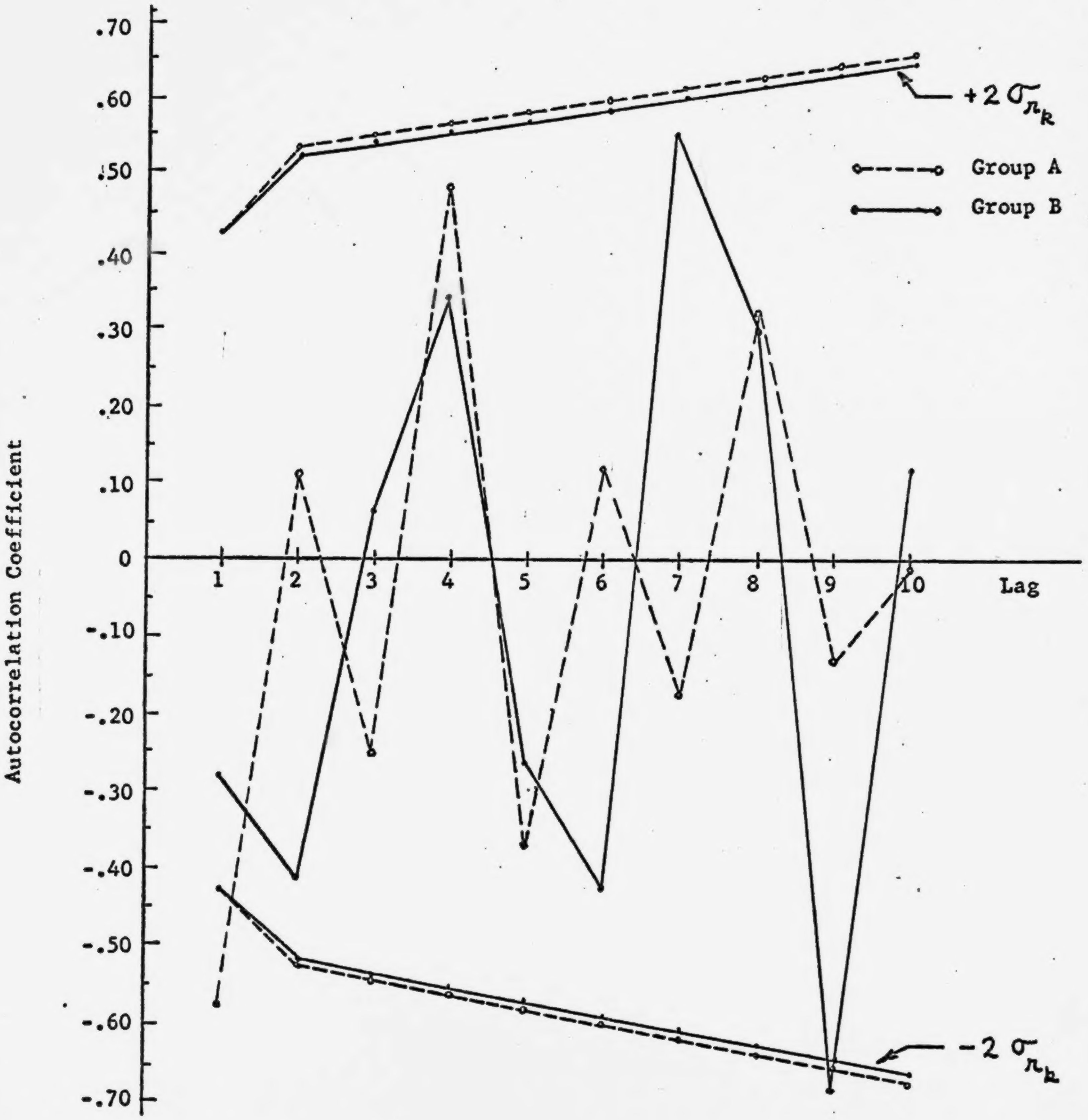
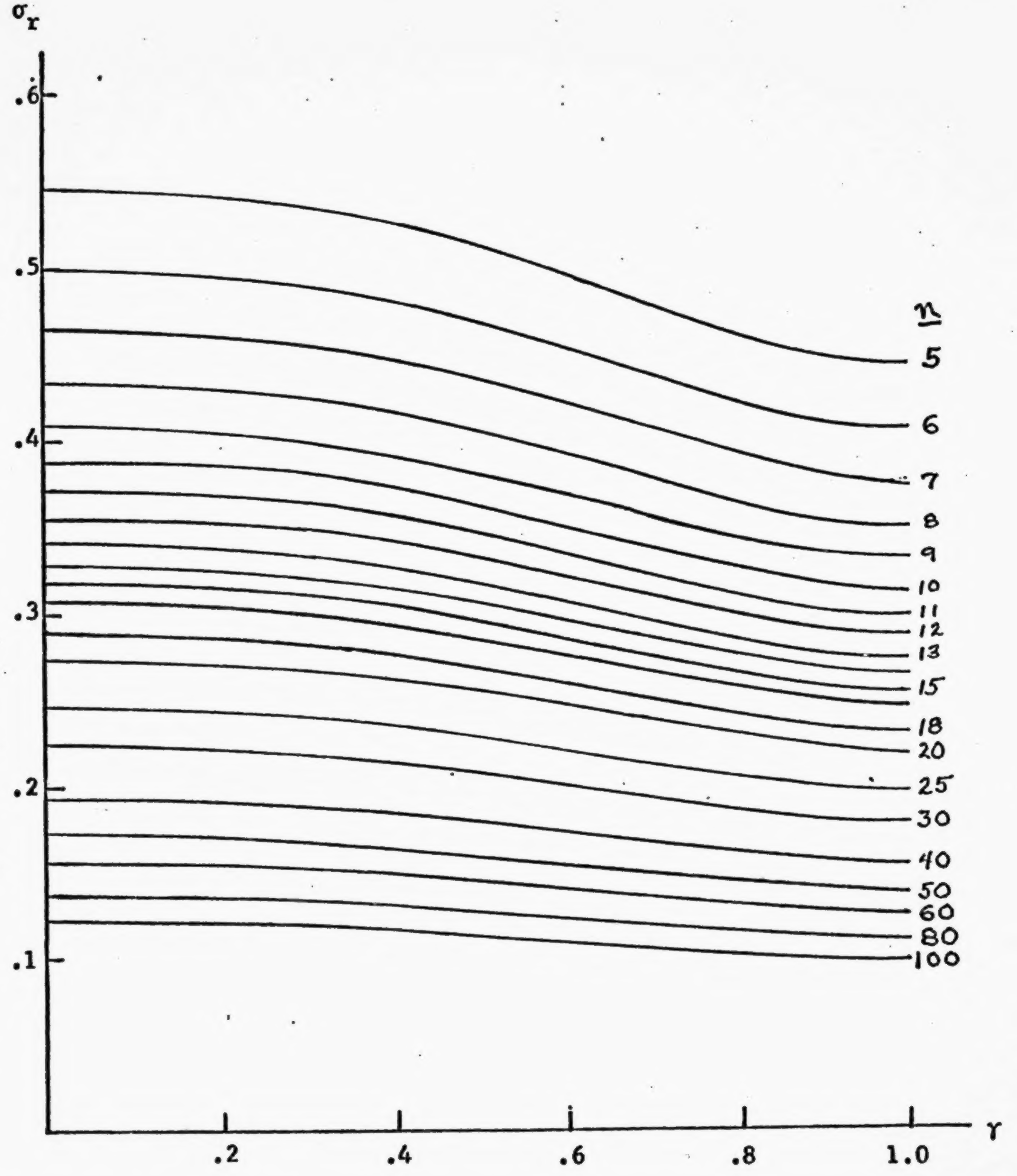


Figure 2. Autocorrelations (lag 1 through lag 10) for differences between adjacent observations in the two time-series in Figure 1.



Approximate standard errors of the lag 2 and greater autocorrelation coefficients for the differences $z_t - z_{t-1}$, where z_t follows the integrated moving average model.

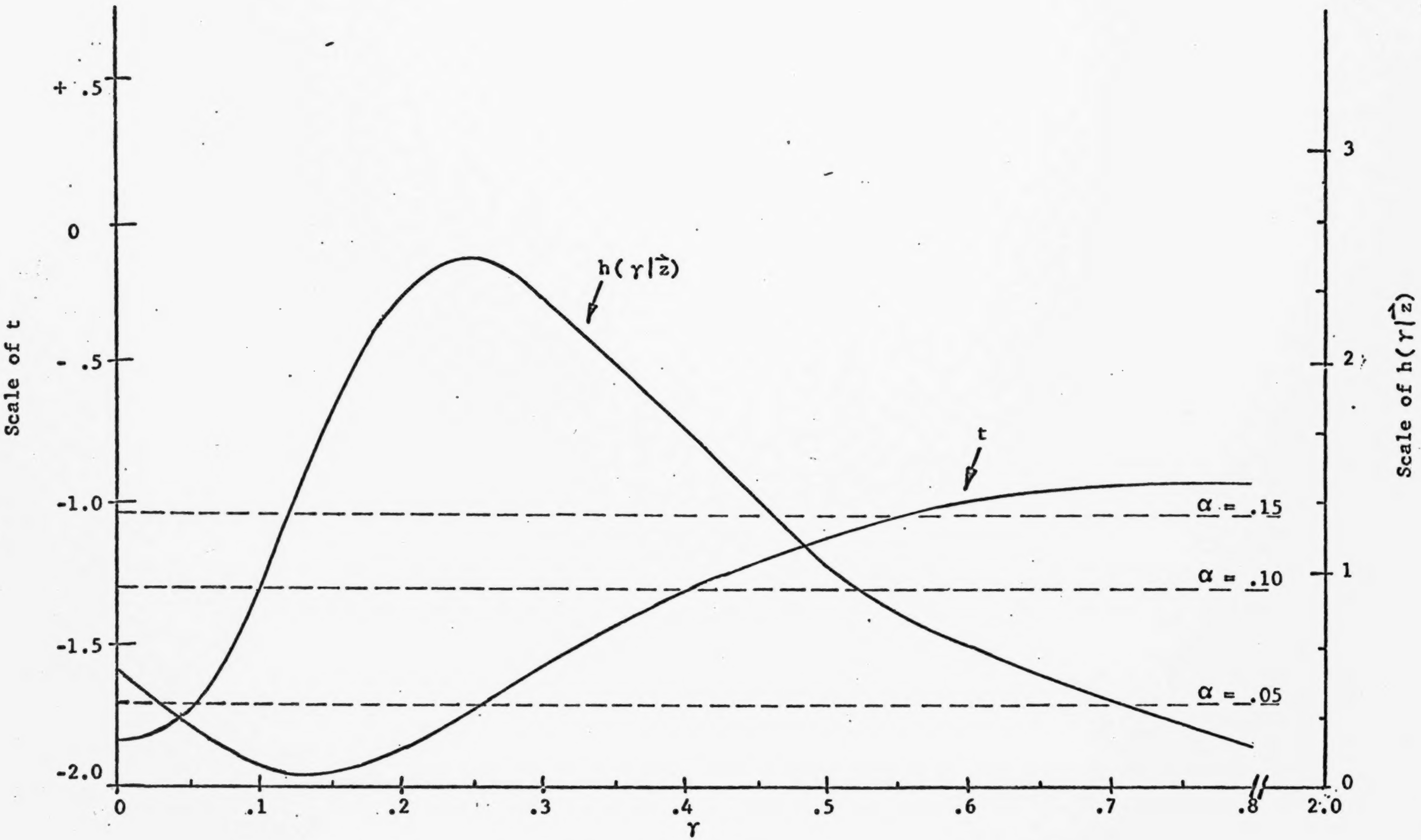


Figure 3. Graphs of the posterior distribution, $h(\gamma|\vec{z})$, of γ and $t = \hat{\delta}/\hat{\sigma}(\delta|\gamma)$ as a function of γ .

(The dotted lines mark the 15th, 10th and 5th percentiles in the t-distribution, dF = 30.)

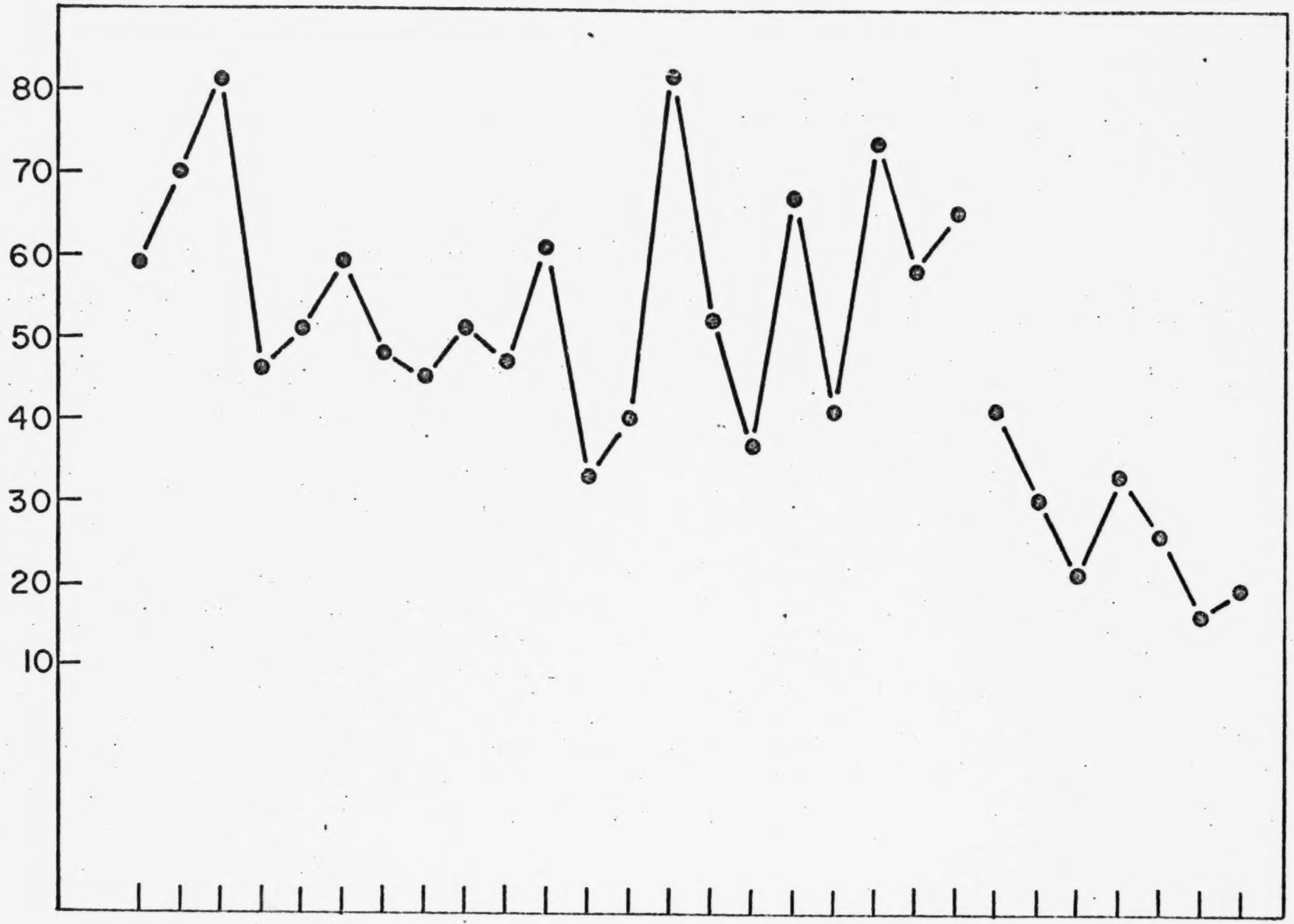


FIG. 4. NUMBER OF ACTIVITY CHANGES IN A 50-MINUTE PERIOD BY S OVER 28 DAYS. SOCIAL REINFORCEMENT GIVEN FOR ATTENDING BEHAVIOR EXCEEDING ONE MINUTE OVER THE LAST SEVEN DAYS.

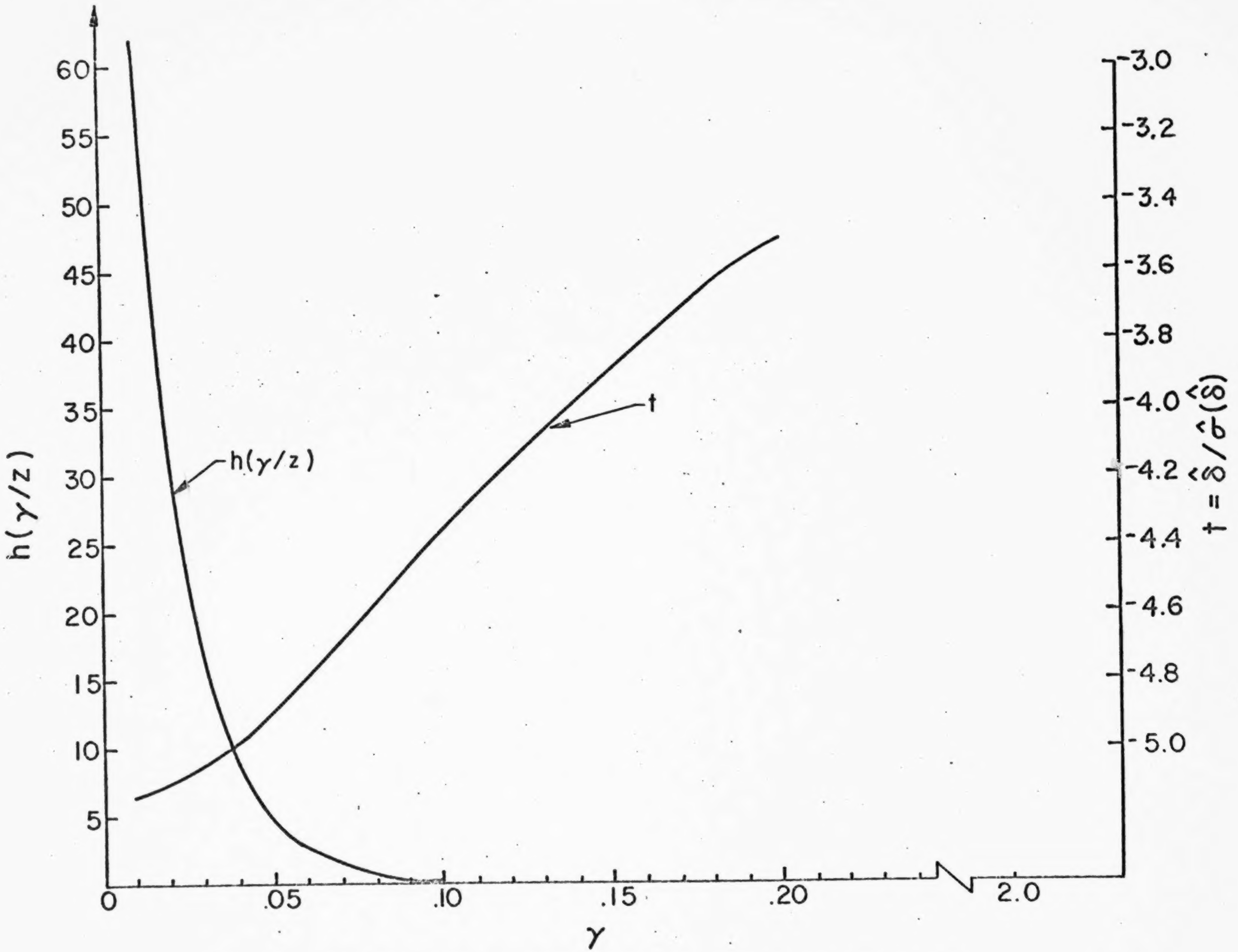


FIG. 5. $h(\gamma/z)$ AND t FOR DATA IN FIG. 4, $(n_1 = 21, n_2 = 7)$.

References

Allen, K. Eileen *et al*. Control of hyperactivity by social reinforcement of attending behavior. *Journal of Educational Psychology*, 1967, *58*, 231-237.

Bartlett, M. S. On the theoretical specification and sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society*, Series B, 1946, *8*, 27.

Box, G. E. P. and Tiao, George C. A change in level of a non-stationary time series. *Biometrika*, 1965, *52*, 181-192.

Brown, Judson S. *The Motivation of Behavior*. New York: McGraw-Hill, 1961.

Campbell, Donald T. and Stanley, Julian C. Experimental and quasi-experimental designs for research on teaching. Chapter 5 in N. L. Gage (Ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.

Chapter III

ANALYSIS OF DATA ON THE CONNECTICUT SPEEDING CRACKDOWN AS A TIME-SERIES QUASI-EXPERIMENT

In late 1955 in Connecticut, the number of fatalities per 100,000 population in motor vehicle accidents reached a record high for the 1950's. On December 23, 1955, Governor Abraham Ribicoff took unprecedented legislative action to reduce traffic fatalities. Ribicoff announced that persons convicted of speeding would have their licenses suspended for thirty days at the first offense, for sixty days at the second offense, and for an indefinite period (subject to a hearing after ninety days) at the third offense. Data on traffic fatalities before and after the Connecticut crackdown on speeding can be regarded as a time-series quasi-experiment (Campbell and Stanley, 1963; Campbell, 1963) with some significance for the social sciences. When supplemented with traffic fatality data for the states of Massachusetts, Rhode Island, New York, and New Jersey, the collection of observations can be viewed as a multiple-group time-series experiment (Campbell and Stanley, 1963; Campbell, 1963). The multiple-group time-series design can be diagrammed as follows:

			Time		
	...$n_1 - 1$	n_1		$n_1 + 1$	$n_1 + 2$...
Conn.:	...0	0	T	0	0...
Mass.:	...0	0		0	0...
R.I. :	...0	0		0	0...
N.Y. :	...0	0		0	0...
N.J. :	...0	0		0	0...

The 0's represent monthly observations of traffic fatalities for the $\underline{n}_1$ months prior to T, the treatment, and for the $\underline{n}_2$ observations following T. The treatment, T, is the Governor's crackdown on speeding in the state of Connecticut. No comparable alteration of the legislation of the four "control" states took place.

Evidence of the effectiveness of the Connecticut crackdown on speeding can be gained by comparing the path of the post-T observations of Connecticut with those of the four control states. A sharp drop in fatalities in Connecticut following T in the absence of similar drops in the control states is compelling evidence of the effectiveness of the crackdown on speeding.

The problem of measuring the abrupt change in level of a time-series and making statistical inferential statements about it is the problem with which the remainder of this report is concerned.

Analysis of Data

The Underlying Model

The statistical model upon which analysis of the Connecticut speeding data is based was developed by Box and Tiao (1965). Box and Tiao presented an analytic technique for estimating and making inferences about the change in level of a non-stationary time-series. The model upon which the analysis is based is a restrictive one; however, many sets of data can be manipulated or transformed into special indices in such a way that the assumptions of the model will be largely met. The statistical model here employed is a special case of the integrated moving average process (Box and Jenkins, 1962):

$$z_t = L + \gamma \sum_{j=1}^{t-1} \alpha_{t-j} + \alpha_t, \quad t = 1, \ldots, n \, . \qquad (1)$$

L is a "location parameter" descriptive of the over-all general level of the series,
γ is a parameter which depends upon the interdependency of the observations in the time-series, and
α_t is an observation of a random normal variable with mean 0 and variance σ^2 .

Formula (1) describes the $\underline{n}_1$ observations taken prior to the introduction of a treatment, e.g., the Connecticut crackdown on speeding. The $\underline{n}_2$ observations following the introduction of the treatment into the time-series differ from (1) only in that a treatment effect, δ, is present.

$$z_t = L + \gamma \sum_{j=1}^{t-1} \alpha_{t-j} + \alpha_t + \delta, \quad t = n_2, \ldots, n_1 + n_2. \qquad (2)$$

The parameter δ is the increment or decrement in the level of the time-series due to the introduction of the treatment. The treatment is assumed to work an immediate and constant effect, δ , upon the time-series.

The fundamental time-series model regards the system as being subjected to periodic random shocks, the α_t, (which have zero mean). Furthermore, a proportion, γ, of each shock is assumed to remain in the system to influence the movement of the system through time.

Hence, the effect of some extraneous, random influence on the system is not immediately dissipated but continues to work a lessened influence on subsequent observations. In a sense, the model describes the path of a point taking a random walk which is imbedded in "noise."

If the value of γ is known, least-squares estimates of L, which is generally of no interest, and δ, which is of primary interest, can be readily obtained (see formulas 3.7 and 3.8 in Box and Tiao, 1965). However, one is unlikely to know the value of γ. Not knowing γ, one may use sample information to determine probable values for γ. Box and Tiao (1965) presented a Bayesian approach to obtaining information about the value of γ. When a uniform prior distribution is set on γ, the Bayesian analysis is equivalent to inspecting the likelihood function of γ. The likelihood of γ as a function of the sample data is given in equation (5.8) in Box and Tiao (1965).

The analytic strategy which the Box-Tiao procedure leads to in most instances is one of calculating the likelihood distribution of γ from the data, finding an estimate of δ and its standard error for all values of γ from 0 to 2.0, and setting confidence intervals around the estimate of δ or testing the significance of the difference of the estimate of δ from zero for the likely values of γ.

However, before such an analysis of data may proceed, an effort must be made to check the appropriateness of the model in (1) and (2) for the data in hand. This can be done in large part by inspecting the graph of the time-series and of the correlograms of the data. Data which conform to the model in (1) and (2) will have the following properties:

1. There will be an absence of cycles in the original data, i.e., the $\underline{n}_1 + \underline{n}_2$ observations z_t. The data will appear to fluctuate around a constant elevation, L, with only minor or momentary drifts away from this baseline. In other words, sustained "drifts" from a baseline in one direction probably indicate a violation of the model, viz., α_t probably has a non-zero mean.

2. The correlogram of the original data, the z_t, is free of cycles and shows a random fluctuation around a baseline. The correlogram does not show the familiar damped cyclic curve characteristic of autoregressive time-series. (In a time-series quasi-experiment it is necessary to calculate correlograms separately for pre-treatment and post-treatment observations, since a large treatment effect will produce strong lag correlations.)

3. The correlogram of the differences between successive observations in the time series, i.e., the correlogram for $z_t - z_{t-1}$ $(t = 2, \ldots, \underline{n}_1 + \underline{n}_2)$ has a lag 1 correlation which is large in absolute value when γ deviates from 1.0 and all higher lag correlations are near zero. In fact, $\text{cov}(z_t - z_{t-1}, z_{t+1} - z_t) = -(1-\gamma)\sigma^2$ and $\text{cov}(z_t - z_{t-1}, z_{t+k} - z_{t+k-1}) = 0$ for $k > 1$. The associated lag 1 autocorrelation of the differences is $\dfrac{-(1-\gamma)}{1 + (1-\gamma)^2}$.

   For example, if $\gamma = 1$, all lag correlations of the differences between successive values are expected to be zero. Fortunately, approximate hypothesis tests are available for testing the significance of the lag correlations (Bartlett, 1946).

Investigation of the Fit of the Model to the Data

The basic data were traffic fatalities for the 60 months prior to the Connecticut speeding crackdown in January 1956 and for the subsequent 40 months for Connecticut, Massachusetts, Rhode Island New York and New Jersey. As the first step in the investigation of the fit of the integrated moving average model to these data, each monthly fatalities count was divided by the number of miles driven in the state during that month. The transformed raw data thus became "monthly fatalities per 100,000,000 miles driven" for all five states. Such a transformation would effectively eliminate any upward linear trend in the data [no such trend may appear in the integrated moving average process in equation (1)] due to increases in population, number of drivers, number of cars, etc.

Inspection of the plot of "monthly fatalities per 100,000,000 driver miles" showed marked yearly cycles, as one might expect. The "peaks" of the cycles coincided with the winter months (Dec. - Feb.); the "valleys" occurred during the summer. Such cycles are a clear violation of the assumptions of the integrated moving average model. The correlogram for "fatalities/100,000,000 miles" for Connecticut showed the "damped sine curve" with a period of 12 months which is characteristic of data possessing yearly cycles. (The manner in which the cycles were removed from the data will be discussed later.) It will be instructive for the moment to observe the "monthly fatalities per 100,000,000 driver miles" with the cycles left in. These data appear in Figure 1.

It can be seen in Figure 1 that the fatalities per 100,000,000 driver miles reached the highest point in the period 1951-1955 in December, 1955. To the extent that this "emergency" prompted Ribicoff's decision to crack down on speeding in late December, 1955, the decline immediately following the crackdown can be partly interpreted as the natural tendency of observations chosen for their extremity to regress toward a central value.

There is a marked decrease in fatalities per 100,000,000 driver miles from December, 1955, to January, 1956. However, there are also decreases in fatalities/100,000,000 mi. in six of the eight possible comparisons of a December with the immediately following January. In fact, the drop in fatalities/100,000,000 miles from December, 1957, to January, 1958, is almost equal to the drop from December, 1955, to January, 1956. A natural drop from any December to the immediately following January in fatalities/100,000,000 miles is quite apparent in Figure 1. Such cycles are also obvious in the graphs of monthly fatalities per 100,000,000 miles in the four "control" states.

The following technique was employed to remove the cycles from the data. Since the cycle had a period of twelve months, the average fatalities/100,000,000 miles for each of the nine Januaries (1951-1959) was subtracted from each January observation. Similarly, observations on each of the other eleven months were deviated around the average (over nine monthly values) fatalities/100,000,000 miles for that month. This was done for each of the five states. (A constant, 2 or 3, was then added to these transformed scores to make them all positive.)

These transformed data showed neither apparent cycles nor upward or downward trends. The data appear in Figures 2-6. In this form, the data appear to satisfy the first condition of the integrated moving average model in equation (1). The next step in the examination of the fit of the integrated moving average model to the data involves the correlograms of the transformed observations and the differences between adjacent observations in the series.

Correlograms were calculated on the data in Figures 2-6 for pre-January 1955 (n_1=60) and post-January 1955 (n_2=48) data separately.

(A marked change in level of a time-series due to a treatment effect would alter the autocorrelations from what they would be in the fundamental process which generates the observations in the time-series; hence, in judging the fit of a model to data from a time-series experiment, correlograms must be calculated separately for pre- and post-treatment observations.) To conserve space, these correlograms are not reproduced here. None of them showed the "damped sine curve" characteristic of autoregressive series. Indeed, each correlogram appeared to be no more than a random array of non-significant autocorrelations characteristic of the correlogram to be expected from data conforming to the integrated moving average model.

The next step in the investigation of the fit of the model in (1) to the data is to calculate the correlogram for the differences between adjacent observations, $z_t - z_{t-1}$. It is necessary to calculate these differences separately for the pre-treatment and post-treatment data. Only the correlograms for the 60 pre-treatment observations for each state are examined here. As was pointed out earlier, if the model in (1) is satisfied, the lag 1 autocorrelation of the differences $z_t - z_{t-1}$ will equal $-(1-\gamma)/[1+(1-\gamma)^2]$, where γ is an unknown parameter in the model, and the lag 2 and greater autocorrelations of the same data will equal zero. Not knowing γ, it is necessary to obtain an estimate of it. Later it will be seen how the likelihood distribution of γ can be found from the N observations, z_t, and the maximum likelihood estimate of γ found therefrom. The maximum likelihood estimates for each of the five states were found to be the following:

State	Maximum Likelihood Estimate of γ	Corresponding Expected Lag 1 Correlation of $z_t - z_{t-1}$
Connecticut	.01	-.50
Massachusetts	.01	-.50
Rhode Island	.01	-.50
New York	.16	-.49
New Jersey	.11	-.49

In light of the above data, the correlograms for the 59 observations of $z_t - z_{t-1}$ for each state should present a lag 1 correlation of approximately -.5 and lag 2 and greater correlations which differ insignificantly from zero. The first such correlogram--for the Connecticut data--appears in Figure 7. The jagged line in Figure 7 is the plot of the lag 1 thru lag 30 autocorrelations for the 59 pre-treatment observations $z_t - z_{t-1}$ for Connecticut. The lag 1 autocorrelation of -.555 agrees quite closely with the expected value of -.50. Superimposed upon the graph of the correlogram are two curved lines indicating those points which lie two standard deviations from the mean, zero, in the distribution of the lag _k_ autocorrelation coefficient for samples of size 59 from a population in which the coefficient is zero (see Bartlett, 1946). Only the lag 1 autocorrelation coefficient is significantly different from zero in Figure 7; hence, the conditions of the model--as reflected in the correlogram of $z_t - z_{t-1}$--appear to be met by the Connecticut data.

The correlograms (lag 1 through lag 20) for Rhode Island, Massachusetts, New York, and New Jersey for the 59 pre-treatment observations $z_t - z_{t-1}$ appear in Figure 8. None of the lag 1 autocorrelation coefficients differs appreciably from the expected values of -.50 and -.49. The lag 2 and greater autocorrelations are distributed around zero with only three coefficients (viz., lags 18 and 19 for Massachusetts and lag 4 for New Jersey) lying further than two standard errors from zero. (The curved lines marking off two standard errors in the distribution of the autocorrelation coefficients which appear in Figure 7 can be applied to the data in Figure 8 as well.) The total import of the data in both Figures 7 and 8 is that the conditions of the integrated moving average model in equation (1) which are reflected in the correlograms of $z_t - z_{t-1}$ are reasonably satisfied by the data for the five states.

After transformation of the data and removal of cycles, the data on fatalities for the five states appear to satisfy all of the conditions of the integrated moving average model in equation (1) reasonably well. We shall proceed with the analyses assuming the data are adequately described by such a model.

Analysis for Change in Level of the Five Time-Series

First, we shall consider in turn the individual analyses for changes in level between the 60th and 61st months of the five time-series in Figures 2-6. The analysis of the Connecticut data (Figure 2) will be considered in detail. Summaries of the analyses will be presented for the other four states. After consideration of the individual analyses, the five sources of data will be combined into a single analysis comparing Connecticut with the "control states."

As was pointed out earlier, the objective of the analysis of a possible change in level of a time-series of the integrated moving average type is to obtain a least-squares estimate of δ in (1) and a distributional statement about the estimate. Provided γ is known, this objective is relatively easily attained. The least-squares estimate of, δ, namely $\hat{\delta}$, and an estimate of its standard error, $\hat{\sigma}(\hat{\delta})$, are given in Box and Tiao (1965) by formulas (3.8) and (3.11), respectively. Under suitable assumptions of normality and independence of errors, $(\hat{\delta} - \delta)/\hat{\sigma}(\hat{\delta})$ has a t-distribution with $n_1 + n_2 - 2$ degrees of freedom. A test of the hypothesis that $\delta = 0$ can be carried out with the test statistic $t = \hat{\delta}/\hat{\sigma}(\hat{\delta})$.

However, if γ is not known (as is generally true), it is necessary to obtain information from the sample of N observations about probable values for γ. Either of two strategies might be followed: 1) find the maximum likelihood estimate of γ and estimate $\hat{\delta}$ and $\hat{\sigma}(\hat{\delta})$ using only that maximum likelihood estimate; 2) plot both the likelihood distribution of γ and the value of $t = \hat{\delta}/\hat{\sigma}(\hat{\delta})$ against the value of γ as it ranges between 0 and 2 and see if the t-statistic is clearly significant or non-significant over the range of probable values for γ. [The likelihood distribution of γ given the N sample observations is found from formula (5.8) in Box and Tiao (1965).] Both strategies will be employed at different points in the analyses to follow.*

A. Analysis for Change in Level of the Connecticut Data (Figure 2).

The $n_1 = 60$ observations preceding the crackdown on speeding in Connecticut and the $n_2 = 48$ post-crackdown observations were subjected to the analysis outlined in Box and Tiao (1965) for unknown γ. The likelihood distribution of γ given the 108 observations is denoted by $h(\gamma | z)$ in Figure 9. The area under the curve $h(\gamma|z)$ is one unit. The maximum likelihood estimate of γ is seen to be 0. The curve denoted by t in Figure 9 is the value of $\hat{\delta}/\hat{\sigma}(\hat{\delta})$--read off the right ordinate in the figure--for each value of γ from 0 to 2.

---

*Calculations performed on computer program described in Chapter II.

As can be seen by inspection of the two curves in Figure 9 almost all the mass of the likelihood distribution of γ lies between 0 and 0.25, the former being the maximum likelihood estimate and the latter being quite unlikely; over this range (0 to 0.25) the value of the t-statistic for testing the hypothesis that $\delta = 0$ (against H_1: $\delta < 0$) ranges from -0.86 to -2.05 (from nonsignificance at the .15 level to significance at the .05 level). Note also that t is significant at the .05 level with a one-tailed test only for γ above .12. If γ is set equal to its maximum likelihood estimate, namely 0, t is nonsignificant even at the .15 level. Inspection of the graphs is facilitated by the dotted lines which mark off the values of t (df = 106) required for significance at the .01, .05, .10 and .15 levels for a one-tailed test of the hypothesis that $\delta = 0$. (For the four control states the alternative hypothesis is that $\delta > 0$.)

The analysis reported in Figure 9 will support neither a confident acceptance nor rejection of H: $\delta = 0$. The analysis proved sensitive to the unknown value of γ. H: $\delta = 0$ can be rejected at the .20 level of significance; but a more cautious decision rule, say $\alpha = .01$, would not lead to rejection of H: $\delta = 0$ for any likely value of γ. A conservative hypothesis tester would probably view the data as providing nor support for the rejection of H: $\delta = 0$.

B. Analysis for Change in Level of the Massachusetts, Rhode Island, New York, and New Jersey Data.

In Figure 10, the likelihood distributions and _t_-statistics for testing H: $\delta = 0$ are presented for the four "control" states. In all analyses, the likely values of the unknown parameter γ fall below .30. The maximum likelihood estimates of γ are .01 for both Massachusetts and Rhode Island. For New York and New Jersey, the maximum likelihood estimates of γ are .16 and .11, respectively.

Considering only the value of _t_ for the maximum likelihood estimates of γ, the Massachusetts data yield the only value of $\hat{\delta}$ which differs significantly ($p < .01$) from zero. The _t_-statistics for Rhode Island, New York, and New Jersey do not attain statistical significance at the .15 level with a directional statistical test.

Considering the value of _t_ over the ranges of likely values of γ, neither the Rhode Island, New York, nor New Jersey data present any evidence for a value of $\hat{\delta}$ significantly different from zero. The results for the Massachusetts data are equivocal. At the maximum likelihood estimate of γ, _t_ is significant at the .01 level. At the point on the γ-scale above which approximately half of the area under $h(\gamma|z)$ lies, _t_ is significant with an α between .01 and .05 with a one-tailed test or .02 and .10 with a two-tailed test. The value of _t_ drops below significance at the .10 level for a one-tailed test (or .20 for a two-tailed test) above the point on the γ-scale above which lies approximately 25% of the area under $h(\gamma|z)$.

None of the analyses of the four control states yields compelling evidence of any abrupt change in fatality rate associated with the events in that state immediately prior to January, 1966. The evidence ranges from definitely not supporting the presence of abrupt change in the case of Rhode Island to slightly equivocal as evidence for an abrupt change in the case of Massachusetts. An alternative analysis exists of the multiple-group time-series experiment in which Connecticut is compared with the four "control" states. The difference between fatalities/100,000,000 miles for Connecticut minus the average fatalities/100,000,000 miles for Massachusetts, Rhode Island, New York, and New Jersey embodies an "experimental and control" comparison. These 108 differences have been calculated and graphed in Figure 11.

One might expect that the analysis of the differences between experimental and control time-series would be somewhat less sensitive than the analysis employed above in which the δ's are estimated separately for each series and then combined in planned or post-hoc comparisons. The process of taking difference compounds residual variability and thereby reduces power. The analysis of the data in Figure 11 bears out this expectation.

The maximum likelihood estimate of γ for the data in Figure 11 is .07 (see Figure 13). Thus one would expect a lag 1 autocorrelation for the differences $z_t - z_{t-1}$ to equal -.50 and the lag 2 and greater autocorrelations to be essentially zero. These conditions are reasonably well met, as can be seen by inspecting the correlogram in Figure 12.

In Figure 13, the likelihood distribution of γ for the 108 observations and the t-statistic for testing the hypothesis that $\delta = 0$ are presented for values of γ between 0 and 2. It is apparent from inspection of the two graphs that $\hat{\delta}$ is not significantly different from zero over the entire range of likely values for γ.

In summary, the five individual analyses gave no convincing evidence for an abrupt change in level associated with the Connecticut speeding crackdown for any of the time-series. The analysis of the differences between the monthly observations for Connecticut minus the average for the four control states showed no evidence of a treatment effect in Connecticut. In the following section, a more powerful analysis of the data is reported.

The Analysis of a Planned Comparison of Connecticut with the Control States.

If the time-series for each state can be regarded as independent of the others, well-known inferential statistical techniques can be employed in making comparisons between Connecticut and the four control states. Accordingly evidence was sought concerning the degree of dependence among the time-series for the different states.

Given the normality assumption of the model in (1) and (2), the independence of the various time-series can be demonstrated if the series show no intercorrelation. To reduce the burden of data analysis without a serious reduction in the sensitivity of the test of the hypothesis of no intercorrelation, data for the first 50 months for Connecticut, Massachusetts, and New York were used. Using "months" as the unit across which correlations were computed, the three intercorrelations of these states were computed for the variable "fatalities/100,000,000 miles minus monthly average." The intercorrelation matrix was as follows:

$$\begin{array}{cccc} & \text{Conn.} & \text{N.Y.} & \text{Mass.} \\ & 1 & -.105 & -.061 \\ R = & -.105 & 1 & -.207 \\ & -.061 & -.207 & 1 \end{array} \qquad (n = 50)$$

A test was made of the hypothesis that the 50 triplets of observations were a random sample from a tri-variate normal distribution in which all intercorrelations are zero (Bartlett, 1950). The test statistic, $-[(n - 1) - (2m + 5)/6] \log_e |R|$, is approximately distributed as a chi-square variable with $m(m - 1)/2$ degrees of freedom, where m is the number of variables. The value of the test statistic for the data in question was 2.877, a value exceeded with probability greater than .30 by a chi-square variable with three degrees of freedom. The three series can probably safely be regarded as independent. (As a general procedure, when γ may depart appreciably from 0, it would be better to intercorrelate the estimated residual errors, the α's, using the maximum likelihood estimates of γ for each series.)

A single planned comparison will serve to evaluate the significance of the change in level of the time-series for Connecticut as compared to the changes or lack thereof in the four control states. This comparison has the following form:

$$\psi = \delta_C - (\delta_M + \delta_{RI} + \delta_{NY} + \delta_{NJ})/4$$

The value of ψ is estimated by replacing the parameters with their least-squares estimates; the variance of the comparison is estimated from the common residual variance for the five states multiplied by $[1^2 + 4(1/4)^2]$. The estimated value of ψ divided by the square root of an estimate of its variance follows Student's t-distribution with 5(106)df when $\psi = 0$. (See Hays, 1963, Chp. 14). However, since the estimated change of level effects and residual variances differ for different values of γ, we shall estimate and test the significance of the comparison for the maximum likelihood estimates of γ for each state and for reasonable upper and lower limits to the value of γ for each state.

The hypothesis to be tested is that $\psi = 0$ against the alternative hypothesis that $\psi < 0$. In words, the hypothesis to be tested is that the change in level of the Connecticut series is no different from the average "change" for the four control states against the alternative that it is less.

Because the γ's are unknown, we shall specify a range between which each γ probably lies as well as the maximum likelihood estimate of each γ. The lower limit to the range for each state will be that value of γ below which approximately 25% of the area under the likelihood distribution of γ lies; the upper limit will mark off approximately the upper 25% of the likelihood distribution. The data for estimating and testing the comparisons appear in Table 1.

Table 1

Values of $\hat{\delta}$ and $\hat{\sigma}(\hat{\delta})$ for the Maximum Likelihood Estimate and Reasonable Upper and Lower Limits of γ for All Five States

State		Max. Likeli. Estimate of γ	Reasonable Upper Limit for γ	Reasonable Lower Limit for γ
1. Conn.	γ	.01	.10	.01
	$\hat{\delta}$	-.152	-.594	-.152
	$\hat{\sigma}(\hat{\delta})$	.176	.391	.176
2. Mass.	γ	.01	.15	.01
	$\hat{\delta}$	.472	.259	.472
	$\hat{\sigma}(\hat{\delta})$	.126	.341	.126
3. R. I.	γ	.01	.10	.01
	$\hat{\delta}$	.079	.326	.079
	$\hat{\sigma}(\hat{\delta})$	.276	.617	.276
4. N. Y.	γ	.16	.25	.10
	$\hat{\delta}$	.275	.247	.337
	$\hat{\sigma}(\hat{\delta})$	.289	.375	.233
5. N. J.	γ	.11	.20	.07
	$\hat{\delta}$	.198	.093	.331
	$\hat{\sigma}(\hat{\delta})$	.292	.391	.236

Table 2

Results of Planned Comparisons of $\hat{\delta}$ for Connecticut with the Average $\hat{\delta}$ for Massachusetts, Rhode Island, New York, and New Jersey

	Max. Likeli. Estimate of γ	Reasonable Upper Limit for γ	Reasonable Lower Limit for γ
$\hat{\psi}$	-.408	-.825	-.457
$\hat{\sigma}_{\hat{\psi}}$	.269	.484	.241
$t = \hat{\psi}/\hat{\sigma}_{\hat{\psi}}$	-1.517	-1.705	-1.896
Prob $\{t_{530} < t\}$	.065	.045	.030

For a given set of five values of γ (one for each state, ψ is estimated by subtracting the average $\hat{\delta}$ for Massachusetts, Rhode Island, New York, and New Jersey from the value of $\hat{\delta}$ for Connecticut. The residual variance, assumed to be equal for all five states, is estimated from the average of the residual variances for all states. The values of $\hat{\psi}$,$\hat{\sigma}_{\hat{\psi}}$ and *t* which correspond to the maximum likelihood estimates of and reasonable upper and lower limits to γ are reported in Table 2. The bottom row of Table 2 is the probability of a Student *t*-variable falling below the value of $\hat{\psi}/\hat{\sigma}_{\hat{\psi}}$. (The probabilities in the last row of Table 2 can be interpreted as the smallest levels of significance for which the planned comparison is significantly different from zero with a one-sided test.)

Conclusion

It can be seen in Table 2 that one may conclude that there is a statistically significant reduction associated with the speeding crackdown in fatalities/100,000,000 driver miles for Connecticut as compared with the four control states.

The above conclusion must not be accepted without due consideration of a source of potential invalidity in the experiment. As Ross and Campbell (1965) pointed out, the fact that Governor Ribicoff was prompted to take action in late 1955 by the alarmingly high fatality rate for that period introduces the possibility of a regression effect from the observations immediately preceding his actions to the observations immediately fdlowing. If one observes a time-series for a period of time and selects that observation which appears quite extreme, subsequent observations are likely to be relatively less extreme. The exact extent of any regression effect in the Connecticut time-series experiment is difficult to estimate.

References

Bartlett, Maurice S. On the theoretical specification and sampling properties of autocorrelated time-studies. Journal of the Royal Statistical Society, B, 1946, 8, 27.

__________________ Tests of significance in factor analysis. British Journal of Psychology, Statistical Section, 3, 1950, 77-85.

Box, G. E. P. and Jenkins, G. M. Some statistical aspects of adaptive optimization and control. Journal of the Royal Statistical Society, B, 1962, 24, 297-343.

Box, G. E. P. and Tiao, G. C. A change in level of a non-stationary time series. Biometrika, 1965, 52, 181-192.

Campbell, Donald T. From description to experimentation: interpreting trends as quasi-experiments. Chapter 12 in C. W. Harris, Problems in Measuring Change. Madison, Wisconsin: University of Wisconsin Press, 1963.

Campbell, Donald T. and Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. Chapter 5 (pp. 171-246) in N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Hays, William L. Statistics for Psychologists. New York: Holt, Rhinehart and Winston, 1963.

Maguire, Thomas O. and Glass, Gene V. A program for the analysis of certain time-series quasi-experiments. Educational and Psychological Measurement, 1967, 27, 743-750.

Ross, H. Laurence and Campbell, Donald T. Time-Series Data in the Quasi-Experimental Analysis of the Connecticut Speeding Crackdown. Unpublished manuscript, 1965. Mimeographed, 22 pp.

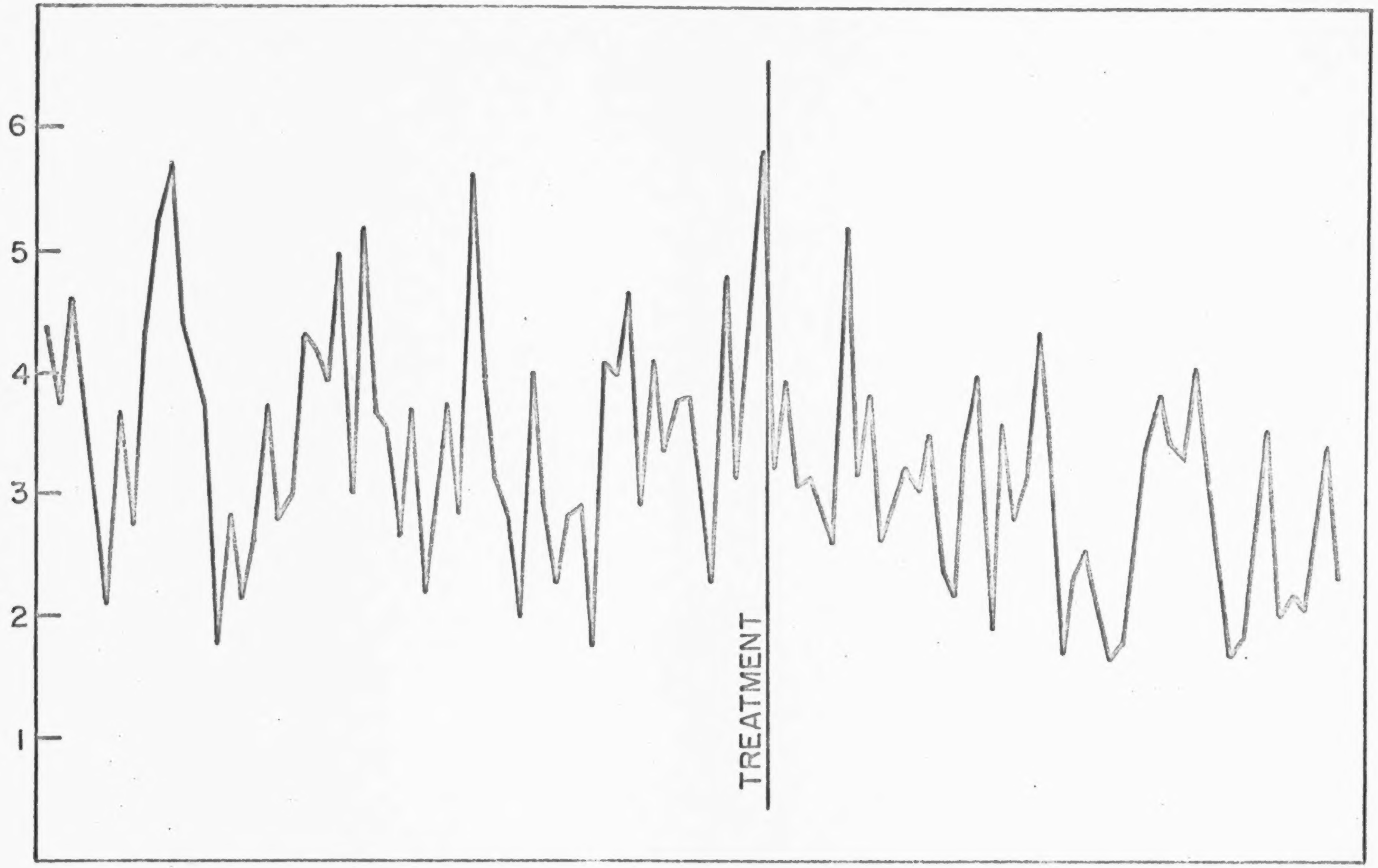


Figure 1. Fatalities/100,000,000 Driver Miles by Months for Connecticut ($\underline{n}_1$=60, $\underline{n}_2$=48).

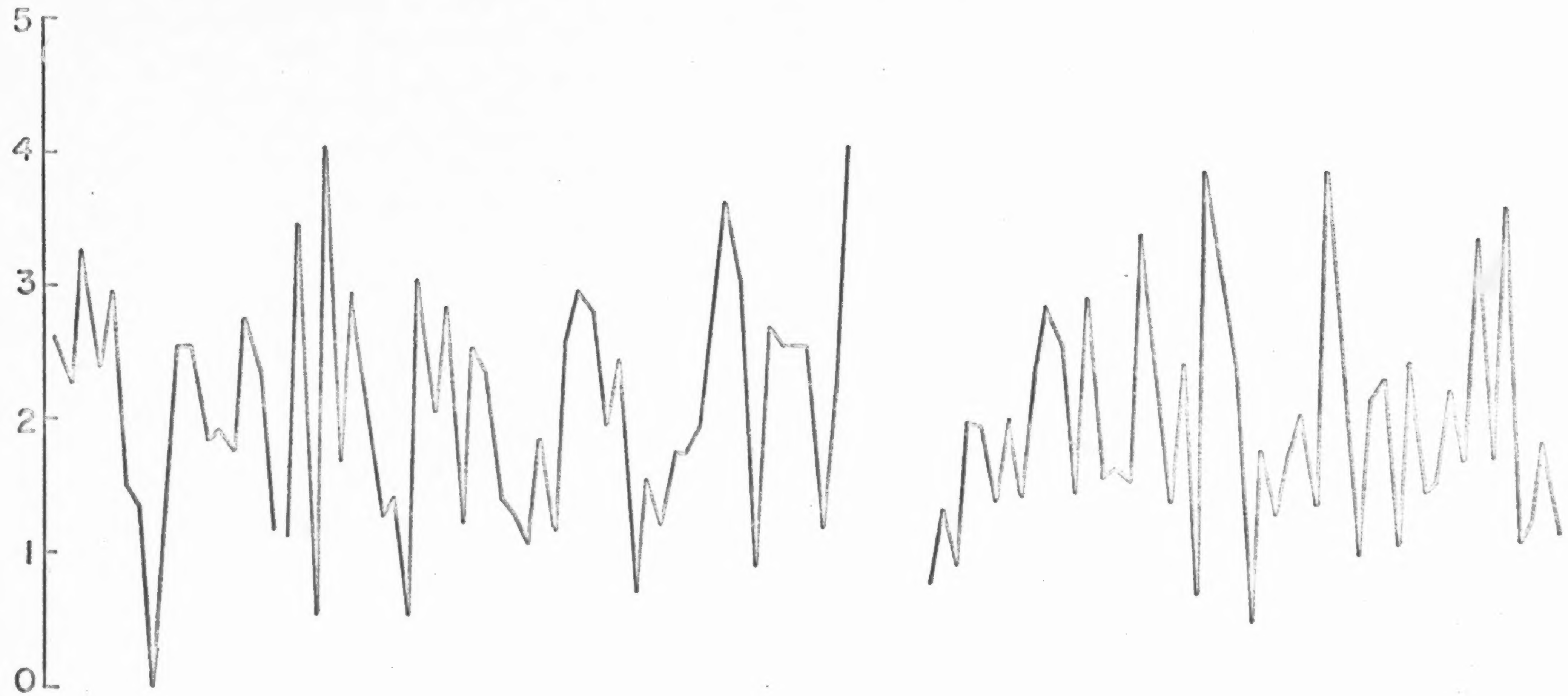


Figure 2. Connecticut Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2.

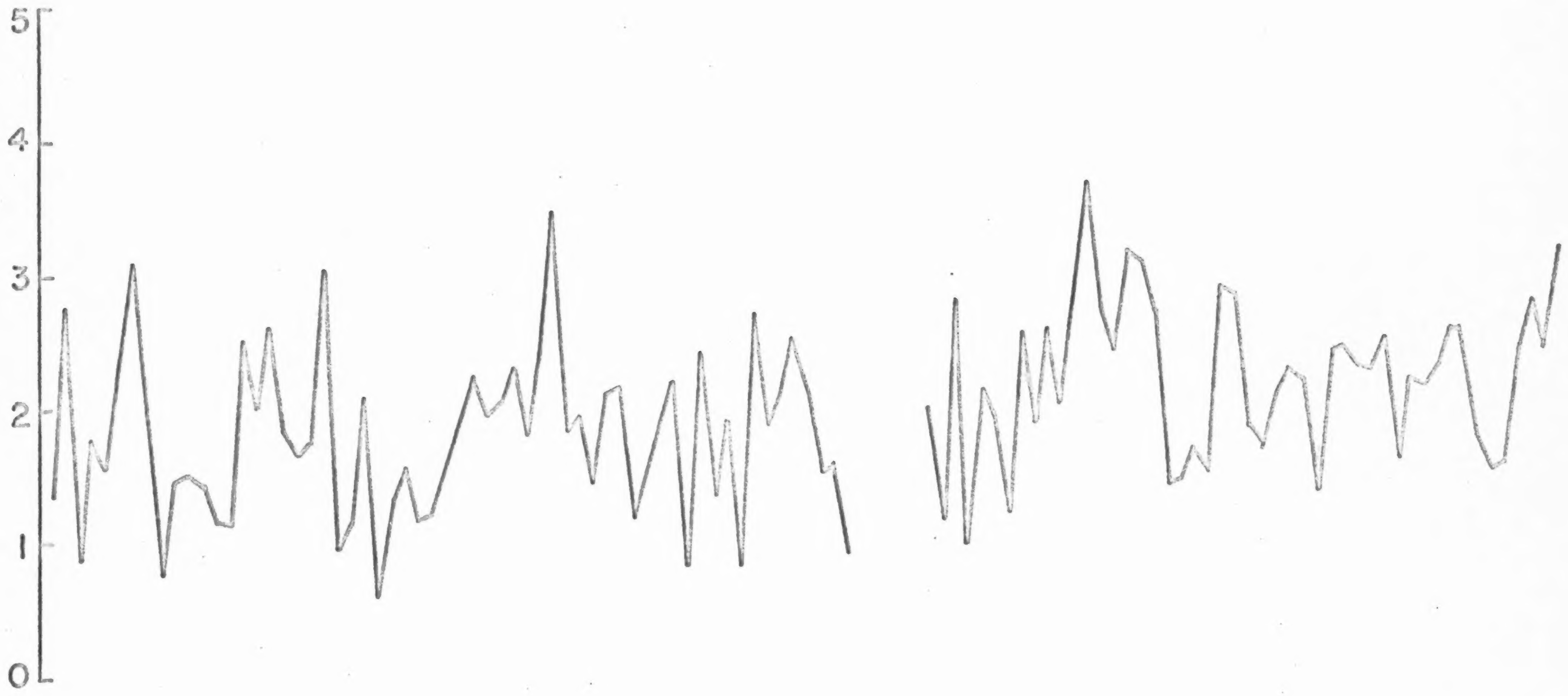


Figure 3. Massachusetts Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2.

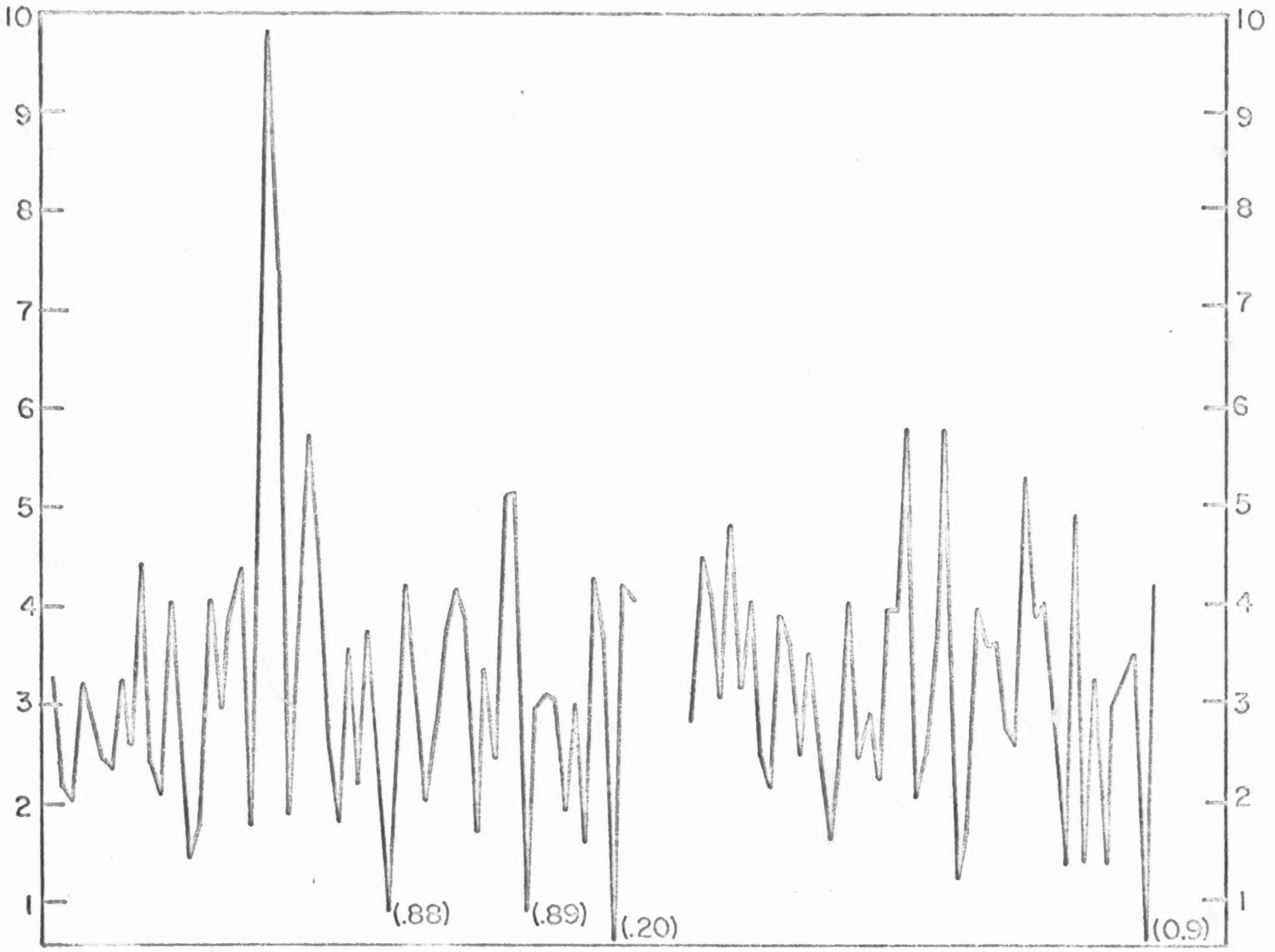


Figure 4. Rhode Island Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 3.

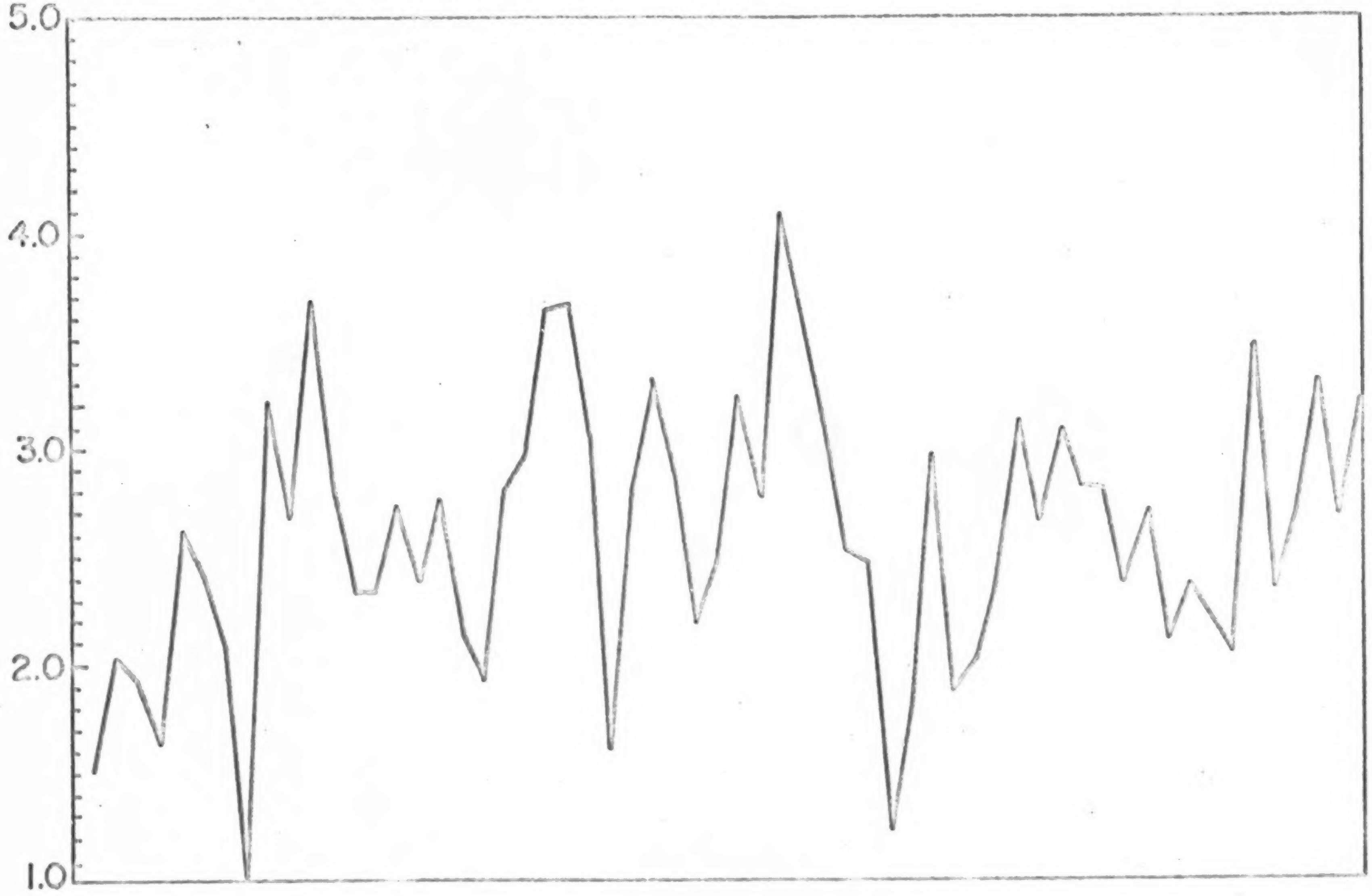


Figure 5. New York Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 3 ($\underline{n}_1$=60, $\underline{n}_2$=48).

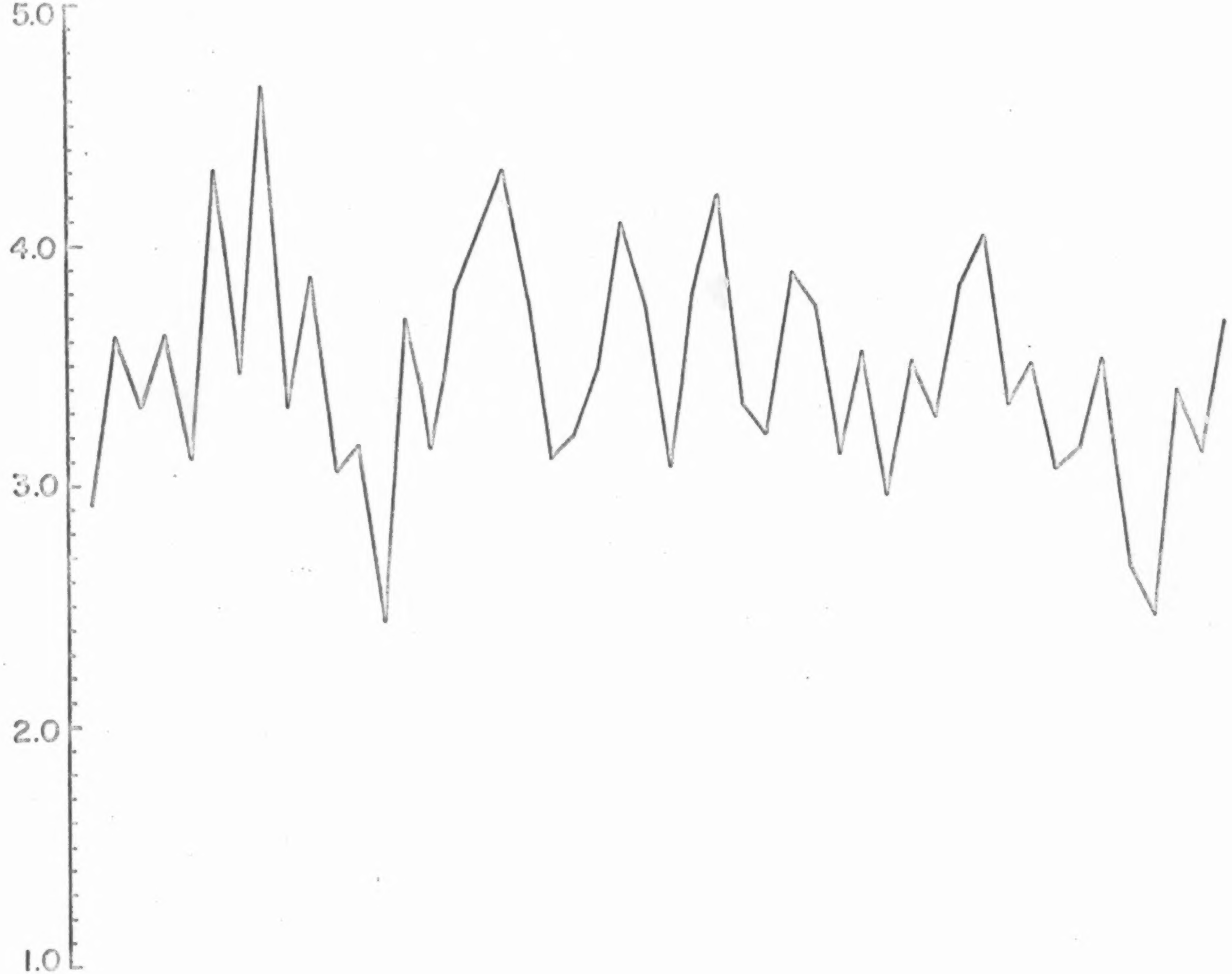


Figure 5. New York Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 3 ($\underline{n}_1$=60, $\underline{n}_2$=48). (Cont.)

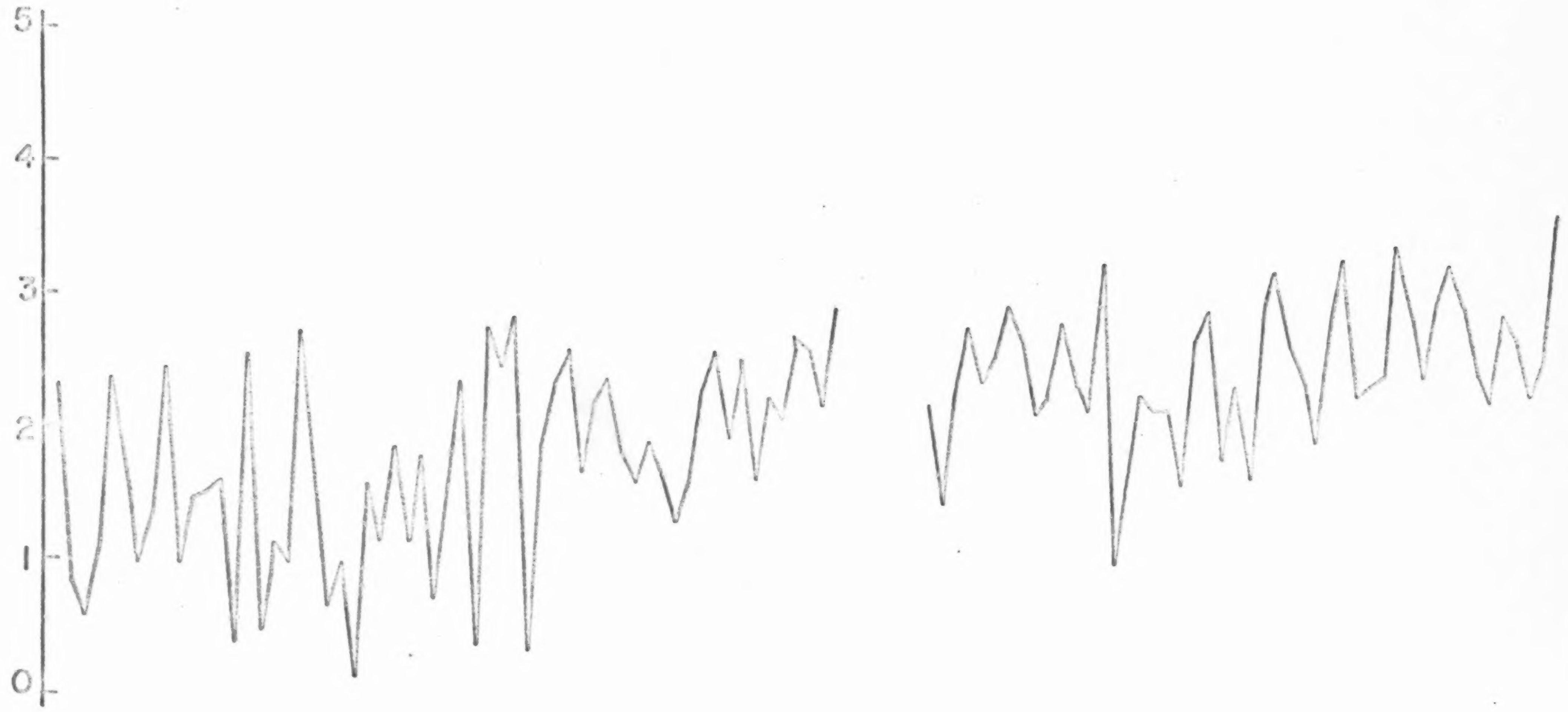


Figure 6. New Jersey Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2.

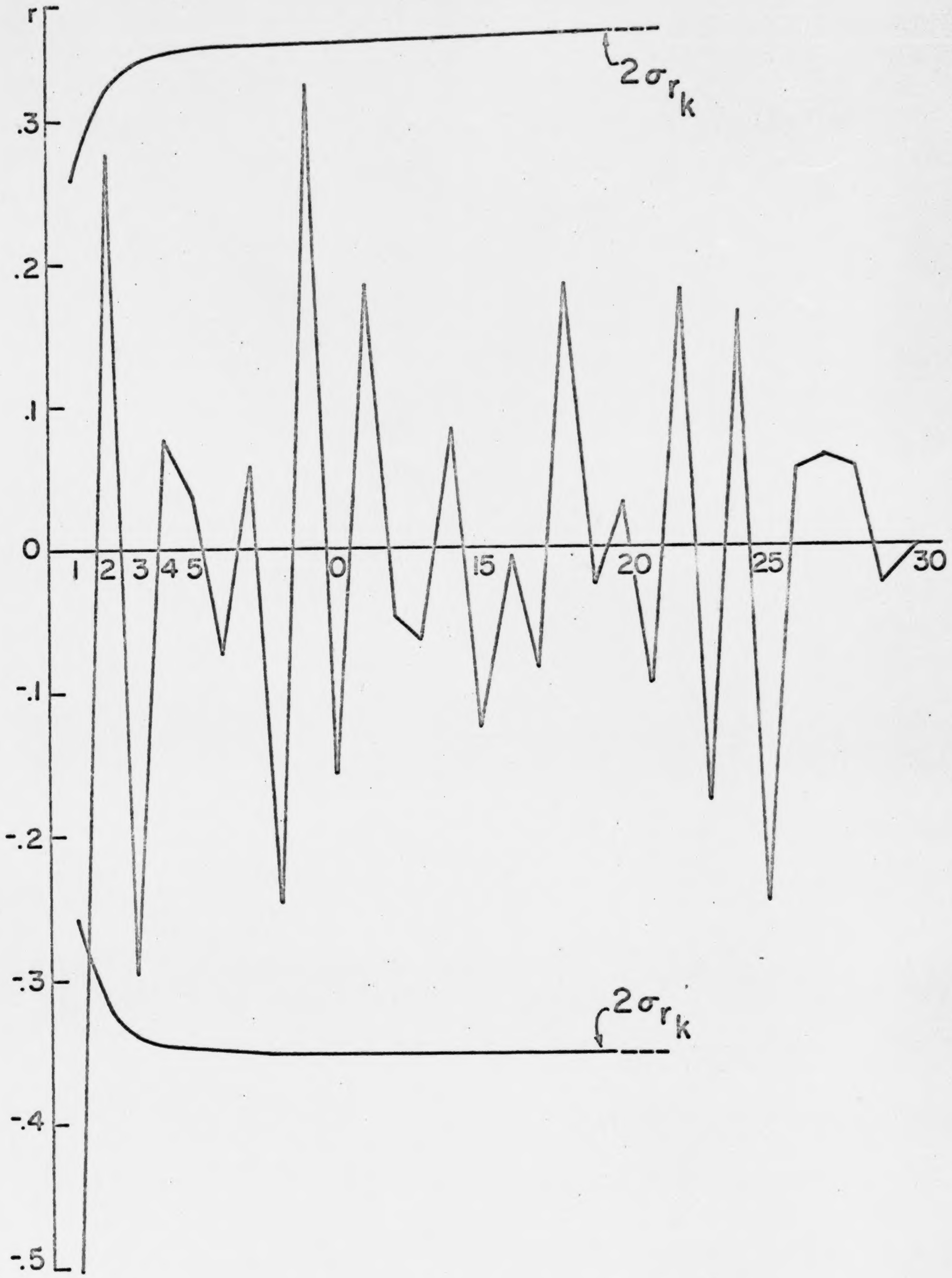


Figure 7. Correlogram of Differences for Pre-Treatment (n_1=60) Monthly Fatalities/100,000,000 Miles Minus Monthly Average Plus 2 for Connecticut.

2

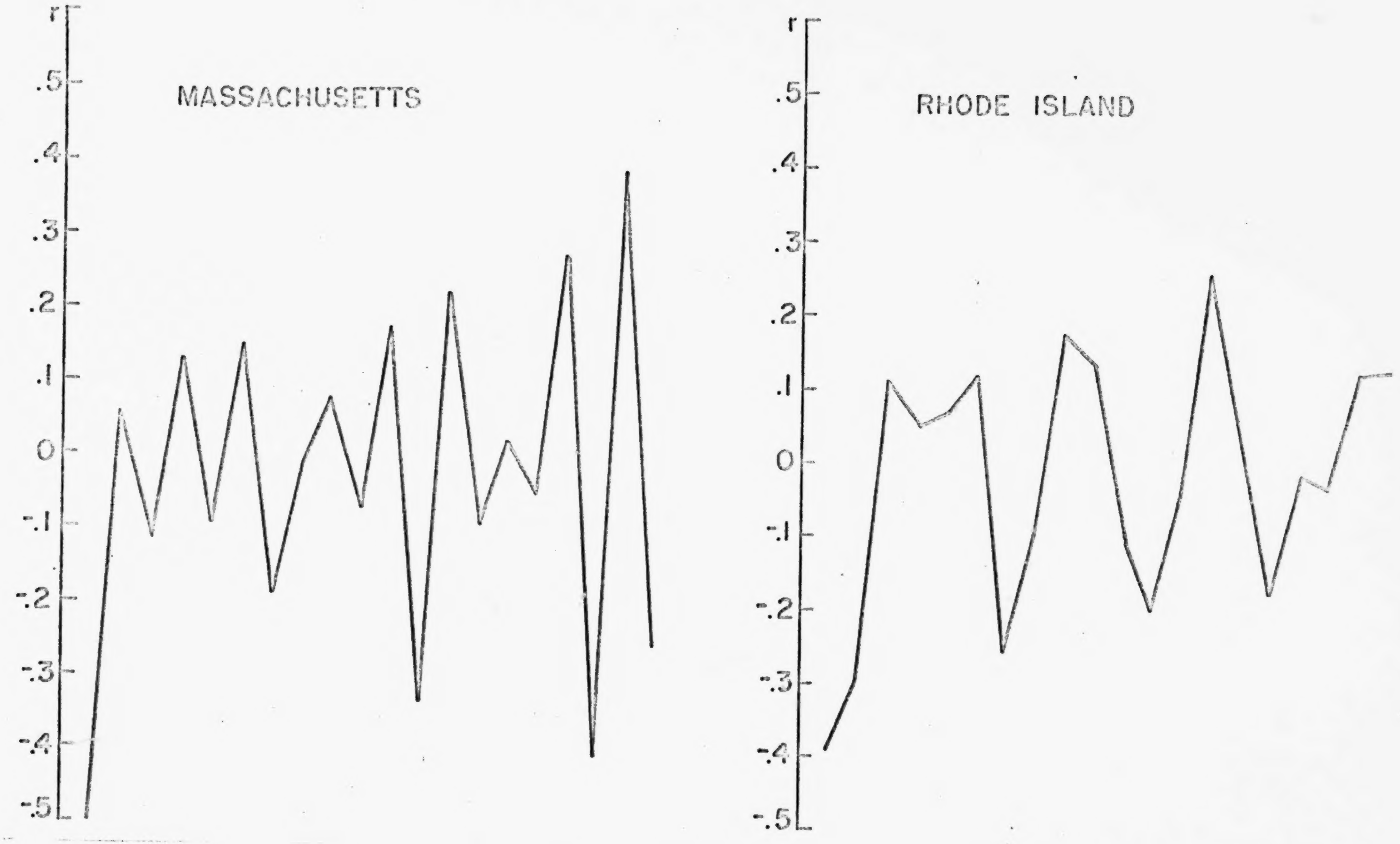


Figure 8. Correlograms (Lag 1 - Lag 20 Autocorrelations) of Differences, $z_t - z_{t-1}$, for Pre-treatment Data ($\underline{n}_1$=60) for Mass., R.I., N.J., N.Y.

(See Figures 3-6 for the data which are "differenced" and correlated.)

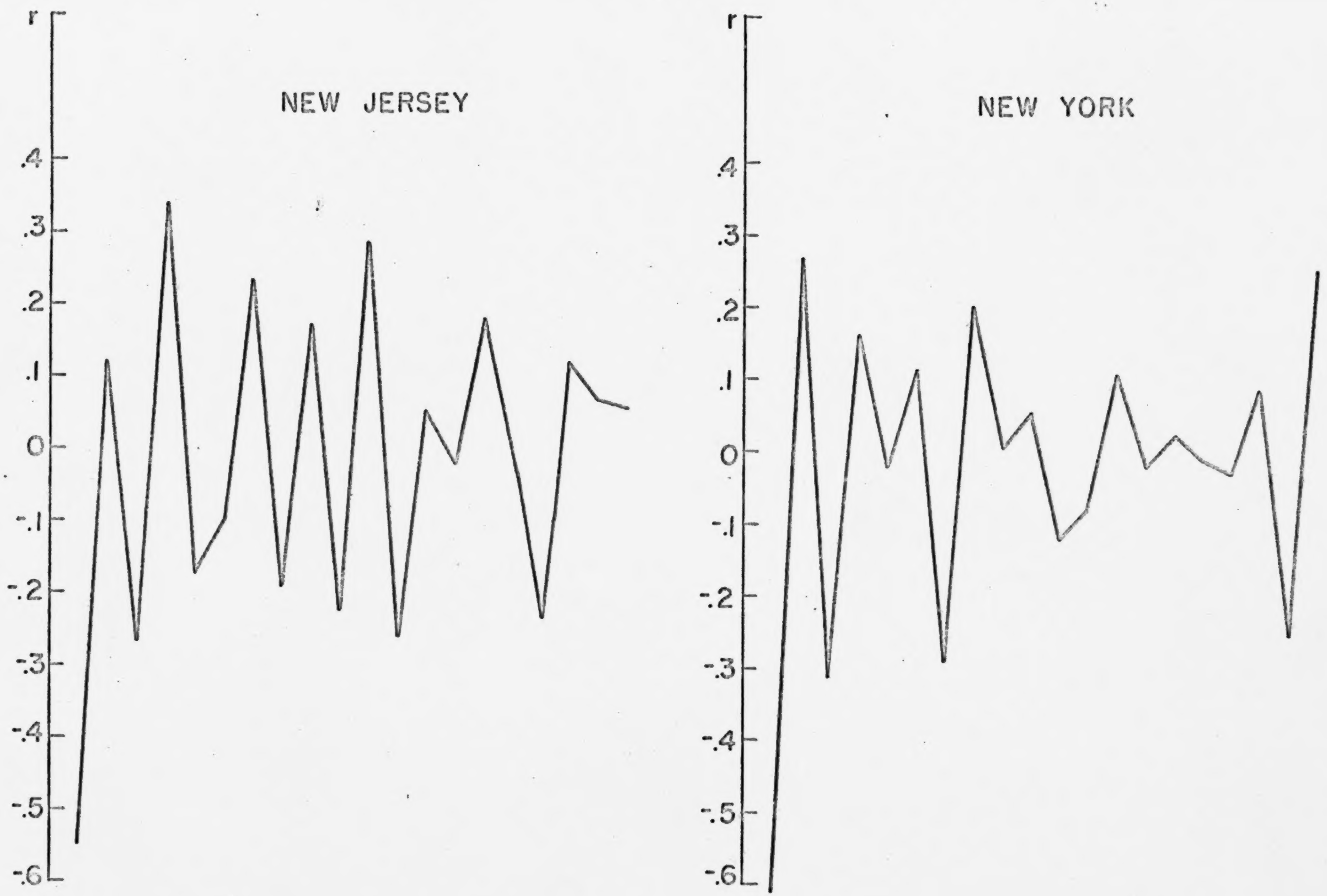


Figure 8. (cont.)

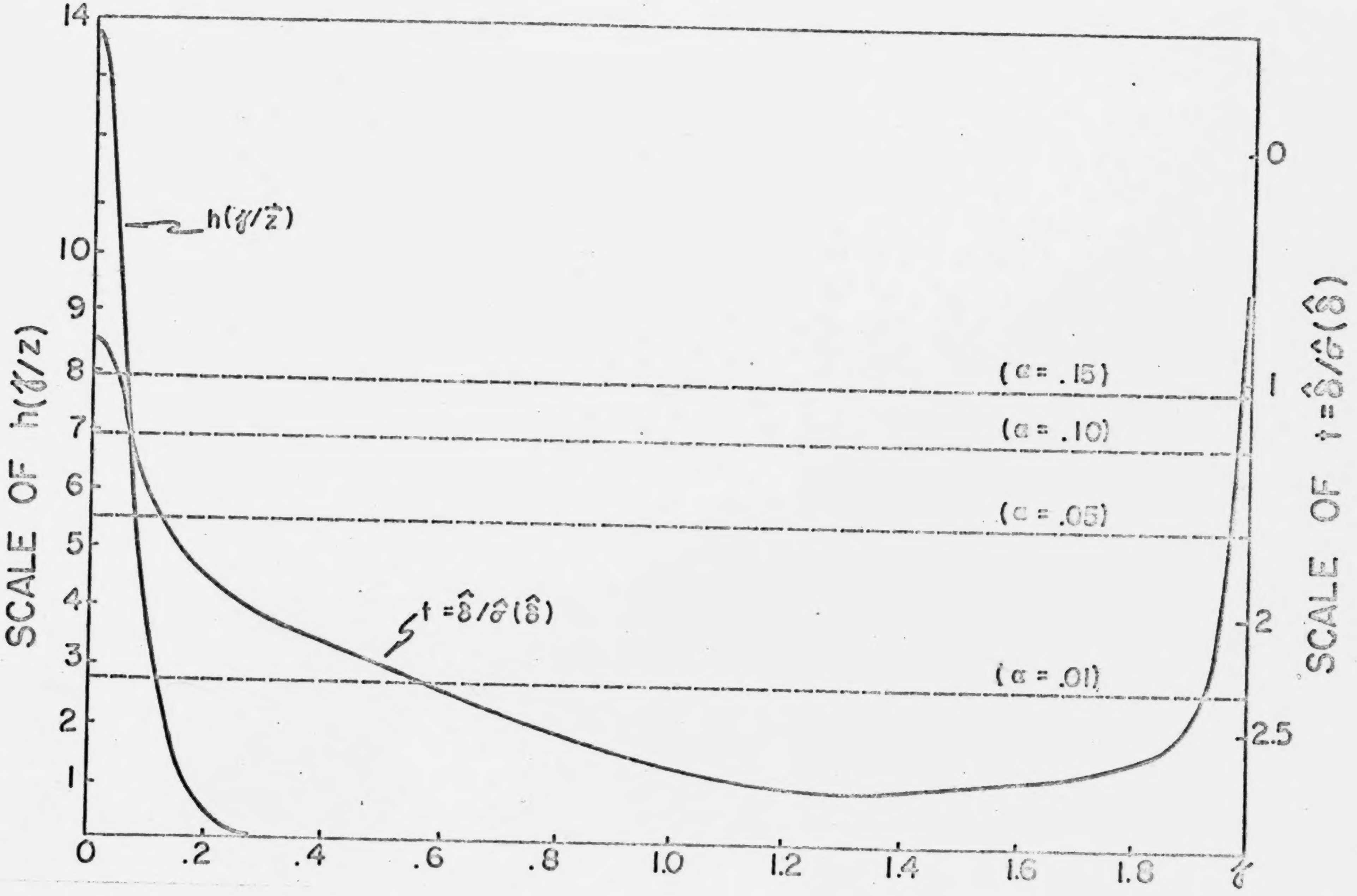


Figure 9. $h(\gamma|z)$ and $\underline{t}$ for Connecticut Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2. ($n_1=60$, $n_2=48$)

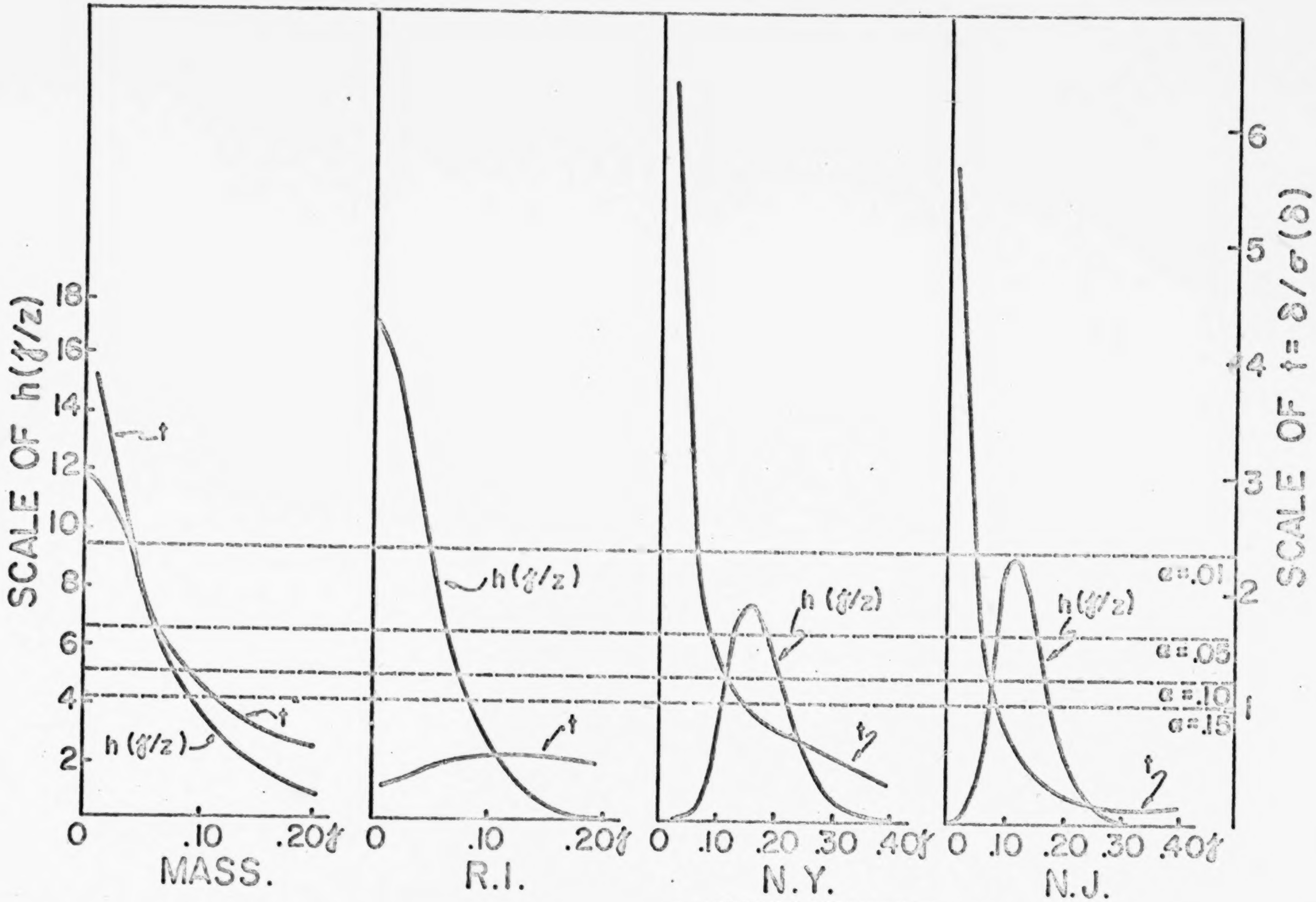


Figure 10. $h(\gamma|z)$ and $\underline{t}$ for Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus a Constant for Massachusetts, Rhode Island, New York, and New Jersey. $(n_1=60,\ n_2=48)$

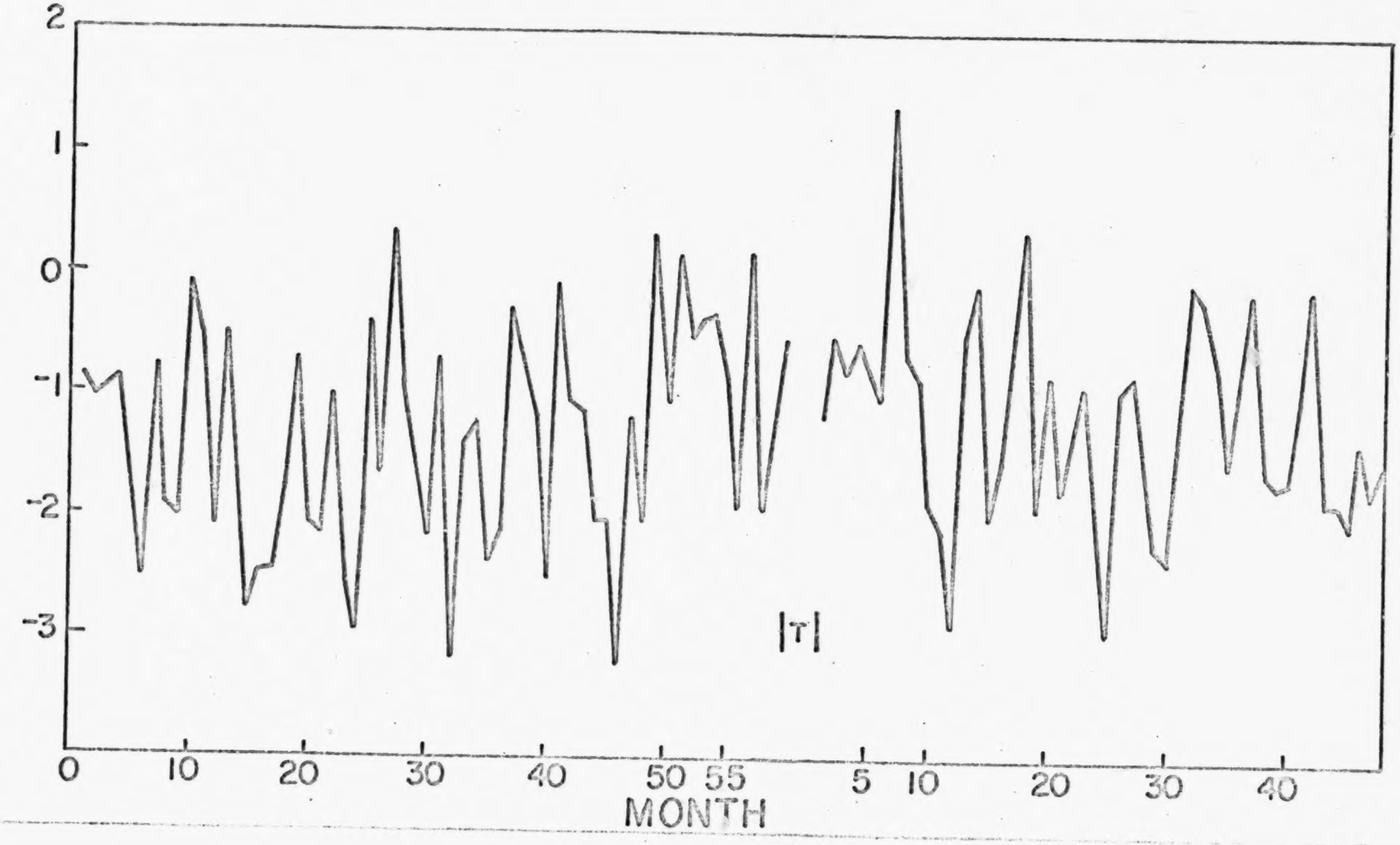


Figure 11. Time-Series of Difference Between Connecticut and the Average for N.Y.-N.J.-Mass.-R.I. of Fatalities Per 100,000,000 Driver Miles by Month ($n_1=60$, $n_2=48$).

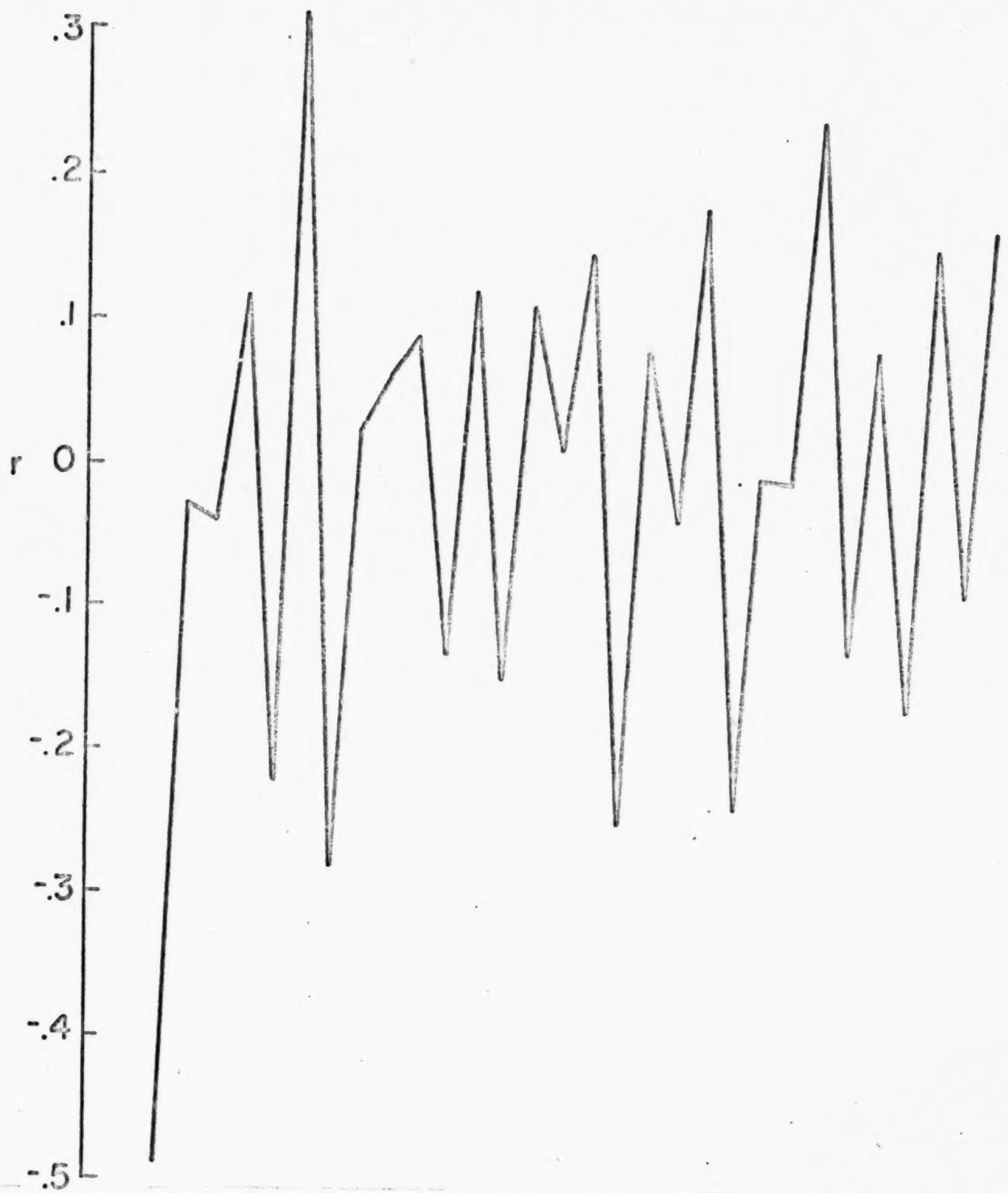


Figure 12. Correlogram of $z_t - z_{t-1}$ for Connecticut Minus Average of Control States

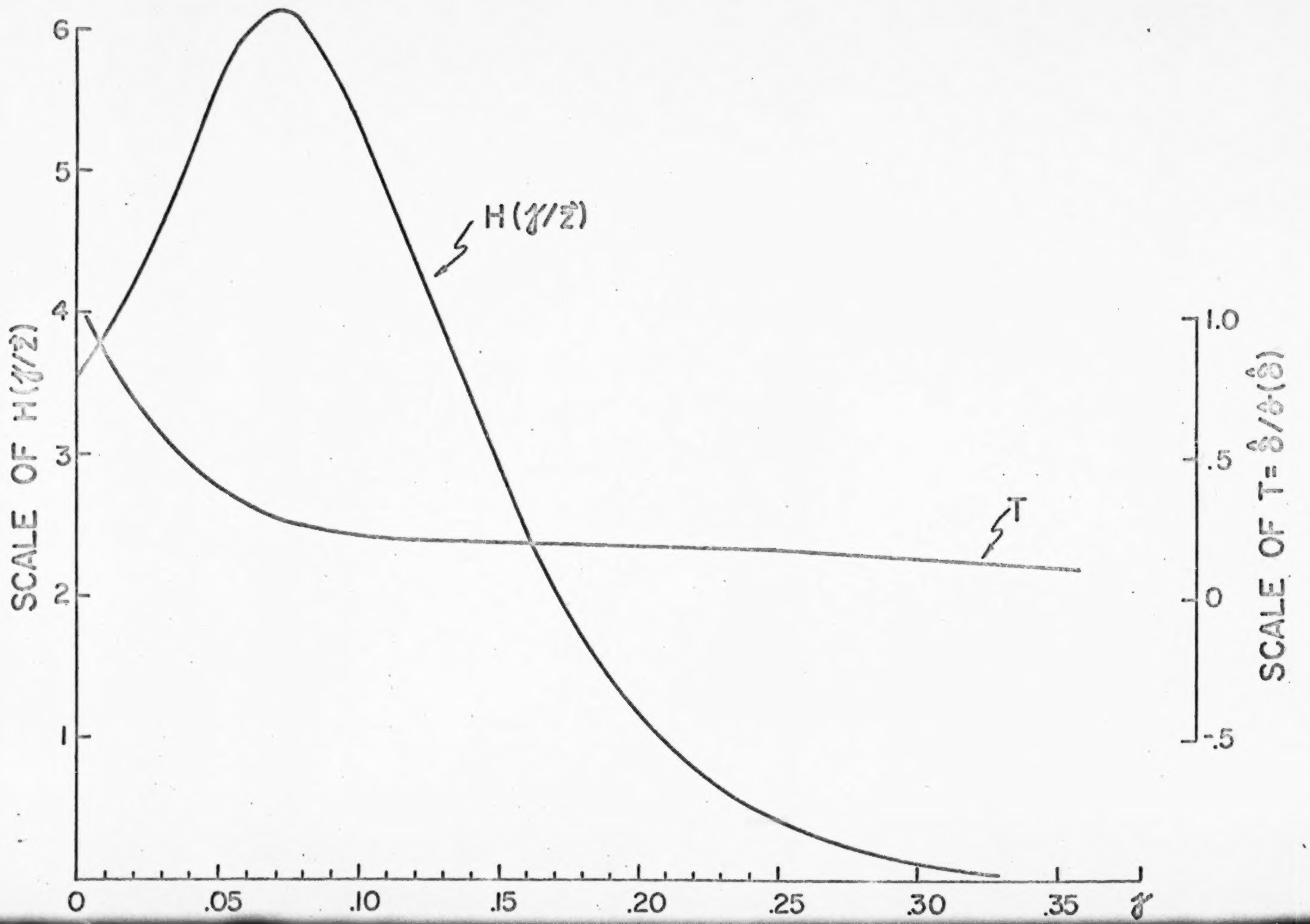


Figure 13. $h(\gamma|z)$ and $\underline{t}$ for Difference Between Connecticut Fatalities per One Hundred Million Miles and the Average of the Fatalities per One Hundred Million Miles for the States of N.Y., N.J., R.I., Mass. by Months.

Chapter IV

Analysis of Data on the Revision of German Divorce Laws as a Time-Series Quasi-Experiment

0. Introduction.

In 1959, Wolf, Lüke and Hax published Scheidung und Scheidungsrecht (Divorce and Divorce Law). This work dealt with the effects of revision of German divorce laws in 1900 on the rates of divorce and petition for divorce. Rheinstein (1959) reviewed their careful study and evaluated their conclusions. The present paper is a re-analysis of Wolf, Lüke and Hax's data with newly developed inferential statistical models and a reappraisal in light of appropriate statistical analyses of the conclusions drawn in both of the above works.

1. The Legislation.

On January 1, 1900, the new Civil Code of the German Empire replaced the various legal statutes then in effect.* The Civil Code

---

*The remainder of this section draws heavily upon Scheidung und Scheidungsrecht by Wolf, Lüke, and Hax (1959) and "Divorce and the Law in Germany: A Review" by Rheinstein (1960).

If the trend of the pre-T observations is altered sharply by the introduction of T, we are inclined to attribute the alteration (whether it be a change in level, change in direction of drift, etc.) to T. A particularly important problem is to determine whether the activity of the time-series in the neighborhood of T indicates a genuine effect of T or whether it is merely an orderly continuation of an undisturbed time-series. We judge the problem to be "particularly important" because the inferential statistical intuitions of social scientists seem seldom to have been developed on non-independent observations (as are in evidence in most time-series), thus formal statistical significance tests are a necessary overseer of "considered impressions" we might form of the data.

The divorce rate (divorce/100,000 persons) for the German Empire from 1881 through 1914 is plotted in Figure 1. In Figure 2 appears the rate of petitioning for reconciliation proceedings for all of Germany from 1881 through 1913. Both indexes are plotted in Figure 3 for those states under the Prussian Code prior to 1900; the same data appear in Figures 4 and 5 for the states under the common law and Code Napoléon prior to 1900, respectively.

3. Statistical Analytic Techniques.

Finding an appropriate inferential statistical analysis of data from a time-series experiment has been repeatedly recognized as an important problem (Campbell, 1963; Campbell and Stanley, 1963). The data in Figure 1 offer an excellent illustration of the need to perform a valid inferential statistical analysis in which the probabilities of

1914. Neither index enjoys unassailable validity as a measure of marital accord. One with faith in the ability of marital partners to repair a disrupted marriage in due time will regard the divorce rate as most significant. To them, the prevention of broken homes at all costs is a worthy goal. Others may regard the rate of petition for reconciliation (which in reality is the initiation of divorce proceedings) as a more valid measure of marital accord; they would argue that marital accord is the more significant variable to attempt to measure since any country can reduce its divorce rate to zero by making divorce illegal (as witness, Italy) without materially affecting the stability of the home. This is not the place to evaluate the social and human value of legal vs. illegal divorce, although it is entirely within the means of the present-day social sciences to do so. Hence, analyses of both indexes will be performed here.

The period from 1881 through 1914 and the intervening revision of the divorce laws can be regarded as an interrupted time-series quasi-experiment (Campbell and Stanley, 1963) for the purpose of assessing the effects of the legislative change. Diagrammatically, the design of the quasi-experiment is as follows:

$$O_1, O_2, \ldots, O_{\underline{n}_1} \quad T \quad O_{\underline{n}_1+1}, \ldots, O_{\underline{n}_1+\underline{n}_2},$$

where O_j represents the $\underline{j}$th successive observation of the divorce rate, say, and

T represents the "treatment" -- in this case, the revision of the divorce laws.

Protestant marriage. German common law was in effect in 12 states prior to 1900.

Similar in practice to the German common law was the Code Napoléon, which was in effect in approximately four states. Divorce was granted only in cases of guilty misconduct; disruption of a marriage constituted insufficient grounds for divorce. (Divorce by mutual agreement was a legal possibility but rarely occurred in practice due to burdensome legal procedures.)

Under the new Civil Code instigated in January of 1900, divorce was to be granted solely on the grounds of guilty misconduct by one partner (adultery, desertion, extreme cruelty, etc.). Divorce by mutual agreement was abolished. The "enlightened disruptive principle" of the Prussian Code was totally displaced by the "guilt principle" in the new Civil Code. Divorce became far more difficult for those who formerly lived under the Prussian Code; it became generally easier to obtain for those formerly under the common law (divorce was legally available to Catholics in ex-common law states for the first time).

2. The Data and the Design.

Two sets of data are available which bear on the question of what effects if any the revision of the divorce laws had: decrees of divorce per 100,000 population (the divorce rate), and petitions for initiation of reconciliation proceedings per 100,000 population (which were mandatory under German law both before and after 1900). These data were reported by Wolf, Lüke, and Hax (1959) for the period 1881-

brought about a general "tightening up" of divorce laws. The Civil Code divorce laws had been drafted in a spirit of hostility toward divorce and with the intention of reversing the steadily increasing divorce rate. (Divorce per 100,000 inhabitants in Germany rose from 8.7 in 1881 to 17.0 in 1899.) Under the new law, divorce was to be granted only in the case of guilty misconduct; divorce was not to be allowed in cases where there was mutual agreement that the marriage should be dissolved or where circumstances had thoroughly disrupted the marriage.

The new Civil Code was uniform across the German states, whereas divorce laws in effect in the various states prior to 1900 were of three general types. The possible effects of the new Civil Code could have depended upon the particular divorce laws in effect before 1900; thus it will be advisable to analyze the effect of the new Civil Code on the divorce rate for three groups of states -- corresponding to the three types of pre-1900 legislation -- as well as for the German Empire as a whole.

Approximately eight states were under the divorce laws of the Prussian General Code prior to 1900. The Prussian Code was the most lenient as regards divorce. Divorces were granted in cases of misconduct, mutual agreement, and even upon grounds of "insuperable aversion" of one party for the other. The Prussian Code recognized "disruption" of the marriage beyond repair as grounds for divorce.

In contrast to the lenient Prussian Code, the German "common law" embodied ecclesiastical law concerning divorce. Catholics could not divorce, and only grave misconduct was grounds for dissolution of a incorrect decisions can be known exactly and controlled. The data in Figure 1 appear to show the expected drop immediately after the change of legislation in 1900.* In fact, the movement of the divorce rate index is larger between 1899 and 1900 than between any other pair of years. However, the 3 point drop between 1899 and 1900 is only 0.2 larger than the 2.8 rise between 1881 and 1882. It would seem incautious, then, to attempt to draw any conclusions by mere inspection of the data or by the application of intuitive judgment.

Wolf, Lüke, and Hax (1959) considered analyzing the data in Figures 1 through 5 by fitting least-squares regression lines (dependent variable-divorce rate; independent variable-year) to the pre-treatment and post-treatment data separately and testing "whether the two lines connect" or whether the datum for 1900 appears to be a simple extrapolation of the pre-treatment regression line. Their suggestion is equivalent to the "Mood-test" suggested by Campbell and Stanley (1963, p. 213). Wolf, Lüke, and Hax recognized the shortcomings of their suggestion and refrained from any inferential statistical analysis. They stated their concerns about statistical procedures for analyzing their data as follows**:

---

*There are 18 observations for the 19 years from 1881 to 1899. Divorce data were not available for the entire German Empire in 1892 and 1893. The observation graphed half-way between 1892 and 1893 is an estimate determined in Wolf, Lüke, and Hax (1959).

**The following passage was rather freely rendered from the original German, but it is substantially correct.

"If the post-treatment regression line connects directly with the pre-treatment regression line, then the change of laws has not brought about a shift of level. If, however, it lies higher or lower, the possibility exists that we are dealing with an effect of the new laws . . . .

"Where in fact a material shift can be established, it is a question of ascertaining whether this shift is to be ascribed to the influence of the change of laws. This can only be accepted when the data show that the shift of level occurred exactly between 1899 and 1900.

"One could suppose that the data are randomly distributed around the regression line. The shift in level between 1899 and 1900 could be regarded as significant under this assumption if the datum for 1899 lay within the chance region surrounding the pre-treatment regression line but the datum for 1900 fell outside this chance region. One could run the alternative test and establish whether the datum for 1900 lies within the chance region of the post-treatment regression line and the datum for 1899 lies outside of the chance region around this regression line.

"If one were to proceed in this way, then one would have to make use of the standard deviation, σ, in ascertaining the limits of the chance region. The standard deviation of a series of values is given by the quadratic mean of all deviations of the individual values from the arithmetic mean. In the case of a regression line, the deviations of the data from the corresponding predicted value take the place of the deviations from the arithmetic mean. If there exist reasons for assuming that the data are distributed as a Gaussian (normal) distribution, then the probabilities would amount to .6827, .9545, and .9973 that a value deviates less than one, two and three standard deviations, respectively, from the mean. For a deviation of a value from the mean of more than three standard deviations it could be assumed rather safely that a special influence instead of a chance fluctuation is being exhibited. The same conclusion would no doubt be clear if the deviation were merely two standard deviations from the regression line.

"It can not be assumed, however, that the process is representative of the present case. Can it be assumed that the fluctuations of the number of divorces are the result of a neutral and unchanging law of a random distribution? Is the number of divorces described as the result of a series of mathematically isolatable factors? Is the chance region into which this number must fall unequivocally determined when these factors remain constant? The answers to these questions have been given in part previously. [See page 12 in Wolf, Lüke, and Hax (1959).] There exists little inducement to assume that the data are distributed

> around the regression line according to a constant mathematical distribution law let alone according to the law for the Gaussian normal distribution. Hence, the calculation of chance regions in this connection appears to be senseless."
>
> (Wolf, Lüke, and Hax, 1959, pp. 129-132.)

Wolf, Lüke, and Hax seem overly concerned about the validity of the assumption of a normal distribution. And in fact, in the passage quoted and on page 12 of their book they express reservations about the validity of any stochastic model as a representation of a social system. They appear to argue that "chance" is an inadequate explanation of social phenomena for which we can find explanations, and they appear to draw some gratuitous connection between the normal distribution and chance phenomena. We can with good success predict and "explain" human stature; the fact that height tends to be normally distributed in adults does not mean that stature is the result of unknown, chance influences.

A valid inferential statistical analysis is available for time-series experiment data, but it is more difficult than fitting and extrapolating least-squares regression lines.

In 1965, Box and Tiao developed a method of evaluating the change in level between two successive points in time of a non-stationary time-series. Observations z_t are taken at equally spaced time intervals and one wishes to make inferences about a possible shift in level of the time-series associated with the occurrence of an event at a particular point in time. If there is an abrupt shift in the level of a time - series between the third and fourth observations, evidence of a treatment effect may exist.

The statistical model underlying the Box-Tiao analysis of change in level of a time-series was the integrated moving average model.

$$z_1 = L + \alpha_1 \text{ and } z_t = L + \gamma \sum_{i=1}^{t-1} \alpha_{t-1} + \alpha_t \tag{1}$$

for the n_1 observations prior to the introduction of T, and

$$z_t = L + \delta + \gamma \sum_{i=1}^{t-1} \alpha_{t-1} + \alpha_i \tag{2}$$

for the $n_2 = N - n_1$ observations following T, where:

z_t is the value of the variable observed at time *t*,

L is a fixed but unknown location parameter,

γ is a parameter descriptive of the degree of interdependence of the observations in the time-series and takes values $0 < \gamma < 2$,

α_t is a random normal deviate with mean 0 and variance σ^2 ,

δ *is the change in level of the time-series caused by T*.

Essentially the model implies that the system is subjected to periodic random shocks, α_t, (with zero mean) a proportion (γ) of which are absorbed into the level of the series. Data which conform to the model in (1) and (2) are such that the graph of the time-series follows an erratic, somewhat random path with slight, *but no systematic drifts, trends, or cycles*. Data which show a systematic increase or decrease over time -- such as population and various growth curves -- violate the assumption of zero mean for the random variable α . For generality, the random variable portion of the model can be allowed to assume an expected value other than zero; thus "drifting" time-series -- those showing a constant rise or fall over time -- can be accommodated. The generalization of the model in (1) and (2) is called the "integrated moving average model with deterministic drift"* and takes the following form:

$$z_1 = L + \beta_1 \quad \text{and} \quad z_t = L + \gamma \sum_{i=1}^{t-1} \beta_{t-1} + \beta_t, \qquad (3)$$

for the $\underline{n}_1$ observations prior to the introduction of T, and

$$z_t = L + \delta + \gamma \sum_{i=1}^{t-1} \beta_{t-1} + \beta_t, \qquad (4)$$

for the $\underline{n}_2 = \underline{N} - \underline{n}_1$ observations following T,

where L, γ and δ are interpreted as in the model in (1) and (2), but now $\underline{\beta}$ is a normal variable with variance σ^2 and mean equal to μ.

The parameter μ describes the rate of ascent or descent of the time-series.

It is illuminating to express β as $\mu + \alpha$ and manipulate (3) into a form similar to (1):

$$z_t = L + \mu\gamma(t-1) + \mu + \gamma \sum_{i=1}^{t-1} \alpha_{t-1} + \alpha_t \qquad (5)$$

One sees by inspection of (5) that the time-series in (3) will be expected to have "drifted" $\mu\gamma t$ units at time $\underline{t}$.

In the setting of the time-series quasi-experiment, interest centers on estimating δ in (4) and testing its significance. The following steps lead to the least-squares estimate of δ and its distribution.

---

*The "integrated moving average model with deterministic drift" was presented by G.E.P. Box and G.M. Jenkins on pp. 33-34 of "Models for Prediction and Control, III. Linear Non-stationary Models," Technical Report No. 79. Madison: Dept. of Statistics, University of Wisconsin, July, 1966.

By setting $y_1 = z_1$, and $y_t = z_t - \gamma \sum_{j=0}^{t-2} (1-\gamma)^j z_{t-1-j}$, the model can be written as $Y = X\,\Theta + e$ where X is defined as an *N* x 3 matrix of weights as follows:

$$\underline{X}^T = \left[\begin{array}{cccc|cccc} 1 & 1 & \cdots & 1 & 1 & & \cdots & 1 \\ 1 & (1-\gamma) & \cdots & (1-\gamma)^{n_1-1} & (1-\gamma)^{n_1} & & \cdots & (1-\gamma)^{N-1} \\ 0 & 0 & \cdots & 0 & 1 & (1-\gamma) & \cdots & (1-\gamma)^{n_2-1} \end{array} \right]$$

Θ is a 3 x 1 vector such that $\Theta^T = (\mu, L, \delta)$; and e is an *N* x 1 vector of random normal deviates, $e^T = (\alpha_1 \ldots \alpha_N)$, the elements of which have mean μ and variance σ^2.

When γ is known, simple least-squares estimates of μ, L and δ can be found from the familiar solution to the least-squares normal equations:

$$\hat{\Theta} = \begin{bmatrix} \hat{\mu} \\ \hat{L} \\ \hat{\delta} \end{bmatrix} = (X^T X)^{-1} X^T Y \tag{6}$$

The least-squares estimates in (6) each have a *t*-distribution with *N* - 3 *df* when divided by appropriate estimates of their standard error. In particular,

$$(\hat{\mu} - \mu)/(s\sqrt{c^{11}}) \sim t_{N-3}, \tag{7}$$

$$(\hat{L} - L)/(s\sqrt{c^{22}}) \sim t_{N-3}, \quad \text{and} \tag{8}$$

$$(\hat{\delta} - \delta)/(s\sqrt{c^{33}}) \sim t_{N-3}, \quad \text{where} \tag{9}$$

$s^2 = (Y^T Y - \Theta^T X^T X \Theta)/(N-3)$ and c^{jj} is the *j*th diagonal element of $(X^T X)^{-1}$.

The above results follow from the linear model $Y = X\,\Theta + e$ in which the errors, *e*, are assumed to be normal, homoscedastic, and independent.

The quantity s^2 is the residual variance, i.e., the variance of $\underline{y}$ after the model $X\hat{\theta}$ is fitted to it.

All of the above operations on the linear model are made for a given value of γ. When γ is unknown (as will generally be true) a Bayesian analysis using sample information about γ is used in making inferences about δ. The posterior distribution, $h(\gamma|z)$, of γ given a set of $\underline{N}$ observations and assuming a uniform prior distribution is known to within a constant of proportionality. The posterior distribution of γ assuming a uniform prior (in which case the posterior distribution is equivalent to the likelihood distribution of γ) is given to within a constant of proportionality by the following formula:

$$h(\gamma|z) \propto |X^T X|^{-1/2} \, s^{-(N-3)} \,. \tag{11}$$

Illustrations of how the posterior distribution of γ in (11) is considered jointly with $\hat{\delta}$ in making inferences about δ for the simple integrating moving average model in (1) appear in Box and Tiao (1965) and Maguire and Glass (1967).

4. Data Analysis and Results. The data in Figure 1 were subjected to the analysis outlined in Section 3.* In Figure 6 appear graphs of the

---

*One condition which data following the model in (4) must satisfy is stated in terms of the correlogram of the differences between successive observations, i.e., $z_{t+1}-z_t$. The lag 1 autocorrelation coefficient should approximate $-(1-\gamma)/[1+(1-\gamma)^2]$, and lag 2 and greater autocorrelation coefficients of $z_{t+1}-z_t$ should approximate zero for the 18 pre-treatment observations, the lag 1 autocorrelation of the 17 differences $z_{t+1}-z_t$ was 0.127, which corresponds reasonably closely to the expected value -- calculated from the maximum likelihood estimate of γ -- of 0.100. The lag 2 through lag 9 autocorrelations were .027, .033, -.419, -.013, .003, -.434, -.504 and .155, respectively.

likelihood distribution of γ and the t-statistic in (7) for testing the significance of the deviation of $\hat{\delta}$ from a hypothesized value of 0. Nearly all of the mass of the likelihood distribution of γ is contained between the values 0.50 and 1.90. The maximum likelihood estimate of the unknown γ is approximately 1.13. The value of $t = \hat{\delta}/\hat{\sigma}(\hat{\delta})$ is clearly significant -- it is never greater than -4.50 -- over the entire range of likely values of γ. The hypothesis H_0: $\delta = 0$ can be confidently rejected in favor of the alternative that $\delta < 0$. Thus we see that the downward shift of the rising divorce rate after 1900 was quite statistically significant; chance can safely be discounted as the explanation of the downward movement of the time-series after 1900. Inspection of Figure 1 seems to indicate that the effect of revision of the divorce laws was temporary. The conclusion that the effect of the change in legislation was temporary depends upon the perhaps gratuitous assumption that the trend from 1881 to 1913 would have been linear (as opposed to curvilinear) in the absence of legislative change.

The results of the analysis of the data in Figure 2 appear as Figure 7. The dotted lines on Figure 7 indicate the values below which t must fall to allow rejection of $\delta = 0$ in favor of $\delta < 0$ at the .05, .025 and .005 levels of significance. The graphs of $h(\gamma|z)$ and t present a picture of somewhat marginal statistical significance. The value of t is significant at the .05 level and beyond for γ above 1.09. The fact that approximately 80% of the likelihood distribution of γ exceeds 1.09 lends support to rejection of $\delta = 0$ in favor of $\delta < 0$ at a respectable level of significance.

The analysis in Figure 7 of the petition for reconciliation rate data in Figure 2 is particularly interesting in that visual inspection of the time-series leaves an impression of no treatment effect which is at variance with the results of the statistical analysis. Wolf, Lüke, and Hax (1959) and Rheinstein (1959) concluded that the revision of the divorce laws in 1900 had no effect on the rate of petition for reconciliation. It is difficult to interpret whether these authors are using the terms "no effect" to mean "no statistically significant effect," "no socially significant effect," or "no permanent effect." In the first sense, one could reasonably take issue with the conclusion of "no treatment effect." It is not our purpose to argue the validity of conclusions of "no effect" in the second and third senses.

In Figure 8 are presented the analyses for change in level of the divorce rate and the petition for reconciliation rate for the German states grouped by type of legislation prior to 1900. These analyses will be summarized below for the three groups of states in turn.

The average divorce rate and petition for reconciliation rate for 12 Prussian Code states are graphed in Figure 3. The graphs of the data create a distinct impression of a strong effect due to the revision of legislation in 1900. The analyses graphed in the left portion of Figure 8 substantiate the statistical significance of the observed downward shifts in the divorce rate and the petition for reconciliation rate. For the divorce rate, the value of $\underline{t} = \hat{\delta}/\hat{\sigma}(\hat{\delta})$ is never greater than -3.90; $\underline{t}$ is approximately -4 at the maximum likelihood estimate of γ. It can be confidently concluded that the divorce rate shifted its level downward at 1900. The petition for reconciliation rate also showed a significant downward shift at 1900; $\underline{t}$ was less than -3 for all likely values of γ, as can be seen in the upper-left portion of Figure 8.

As was pointed out earlier, the new Civil Code instigated in 1900 constituted a "tightening" of divorce laws in those states previously under the Prussian Code. Introduction of the new legislation should have worked a negative effect upon the divorce and petition for reconciliation rates. Such effects are reliably observable in the data.

The average divorce and petition for reconciliation rates for eight Common Law states are graphed in Figure 4. Inspection of the behavior of both time-series in the vicinity of 1900 would probably lead to no confident conclusions about the possibility of treatment effects. The petition for reconciliation rate increases from 1899 to 1900, but not dramatically so. The decrement in the divorce rate from 1899 to 1900 is even less dramatic, and can not be confidently ruled out as a chance occurrence by mere inspection. The analyses for change in level of the petition for reconciliation and divorce rates at 1900 appear in the middle-upper and -lower portions of Figure 8. In the lower-middle portion of the figure, the graphs of $\underline{h}(\gamma|z)$ and $\underline{t}$ appear for the divorce rate. The value of $\underline{t}$ is less than -3 for all likely values of γ. Hence, the rather small downward shift in the divorce rate is nonetheless statistically significant and can not be reasonably attributed to chance. The shift in level of the petition for reconciliation rate is equally statistically significant for the Common Law states; however, whereas there was a decrement in the divorce rate at 1900, there was a statistically significant increment in the petition for reconciliation rate. It should be recalled that under the new Civil Code, divorce became legal for Catholics in Common Law states for the first time. One might speculate that the data support the conclusion that the new Civil Code brought about petitions for reconciliation from Catholics in the Common Law states, but that the courts held to their newly enacted unsympathetic attitude toward divorce. Of course, such speculation goes far beyond the data.

The average divorce and petition for reconciliation rates for four states under the Code Napoléon prior to 1900 are graphed in Figure 5. There appears to be a downward shift of level in the divorce rate at 1900; however, the petition for reconciliation rate does not appear to have been affected by the introduction of the new Civil Code. This latter observation was borne out by the failure of $\underline{t} = \hat{\delta}/\hat{\sigma}(\hat{\delta})$ to attain significance in the test for a change in level at 1900 of the petition for reconciliation rate. As can be seen in the upper-right portion of Figure 8, $\underline{t}$ falls below +2 for approximately the upper 75% of the likelihood distribution of γ; at the maximum likelihood estimate of γ, namely 0.67, the value of $\underline{t}$ is +1.10. The data on the divorce rate for the Code Napoléon states show a statistically significant downward shift at 1900 (see the lower-right portion of Figure 8).

Conclusions

The conclusions we shall draw from the above analyses will be at variance with those drawn by Wolf, Lüke, and Hax (1959) and Rheinstein (1959). With respect to petitions for reconciliation proceedings, Wolf, Lüke and Hax concluded the following:

> The introduction of the new Civil Code [in 1900] has not reduced the increase of the number of petitions for conciliation proceedings and has thus not reduced the extent of the divorce desire. Preponderantly the new law has not had any effect in this respect. In some regions in which the divorce law was liberalized one can observe a certain increase of the trend. It is by no means certain, however, whether this increase would not have occurred independent of the change in the law. Nowhere was the progressive trend retarded. Even in the regions of the Prussian law, where the divorce law was tightened, the trend did not change in any significant way.*

Rheinstein (1959, p. 493) observed that Wolf, Lüke and Hax were "certainly justified in concluding that the draftsmen of the new code have failed in their expectation of reducing the desire for divorce."

With respect to the divorce rate, Wolf, Lüke, and Hax concluded that "the shape of the law of divorce was neither the cause of the divorce wave nor even one of its essential conditions. In the face of other circumstances, the influence of the law did not make itself felt at all."

Rheinstein (1959, p. 495) concurred:

> . . . before 1900 the [divorce rate] was rising in the districts of most and, since 1900, in those of all appellate courts. In a few court districts the trend shows a slight downward break in 1900. The majority of the latter districts belongs to the region of the Prussian Code, but there are among them also two districts of Protestant common law. In all these districts the

---

*The translation is due to Rheinstein (1959, p. 493).

> break is small, and the trend rose continuously after 1900. While the break in the Prussian law districts may be attributed to the change of the law, _it was insignificant and without lasting effect_. Nowhere did the change turn the trend downward; and nowhere did it prevent its continuous rise. (Italics added.)

Both _Scheidung und Scheidungsrecht_ and Rheinstein's review leave us with the conclusion that

> The experiment made by makers of the Civil Code refutes the notice [sic] that a limitation of the statutory catalogues of grounds for divorce to situations of guilt could result in a reduction of the number of divorces or even in their rate of increase. On the other hand, the present Marriage Law [of 1938] has refuted the apprehension that the introduction of the disruption principle would naturally result in an increase of divorce. _No causal or even statistical connection exists in one direction or the other_.

Rheinstein saw Wolf, Lüke and Hax's work as confirmation of Willcox's conclusion that "the immediate, direct and measurable influence of legislation is subsidiary, unimportant, almost imperceptible." (Willcox, 1897).

We contend that the conclusions just stated make an unfortunate use of the word "significant" and that they depend for their validity upon extrapolations of pre-1900 trends for which there exist as a basis neither compelling logical reasons nor convincing empirical evidence. Furthermore, we feel that the only conclusion which may be drawn from the data with confidence is that the effect of the introduction of the new Civil Code in 1900 is clearly reflected in both the divorce rate and the petition for reconciliation rate.

All too frequently, social scientists extended the meaning of the term "significant" beyond its strictly appropriate sense as it applies to statistical hypothesis testing and made unwarranted interpretations of social value, merit, or importance of data when they are merely inferentially reliable -- the appropriate meaning of "statistically significant." Having been disabused of this confusion in our enlightened age, social scientists now react quite cautiously to the words "significant" or "insignificant" when they are applied to data; they are careful to read inferential reliability into the word and nothing else. Thus, Rheinstein risked serious misinterpretation of the facts when he chose to call the break in the divorce rate curve at 1900 "insignificant" without the benefit of a valid statistical analysis and without apprising the reader of the value system against which he judged the downward shift to be without social value or importance -- the popular sense of "insignificant." We have shown that the changes in level of the divorce and petition for reconciliation rates around 1900 are statistically significant (with the exception of the petition for reconciliation rate in the Code Napoléon states). It does not seem justifiable to refer to the shifts in level as "insignificant" in any inferential statistical sense.

It was also concluded above that if any effect of the 1900 revision of the divorce laws did occur it was "temporary" or "without lasting effect." It was claimed that granting a remote possibility of an effect of the new Civil Code the graphs of the divorce rate and the petition for reconciliation rate quickly returned to a trend line one could extrapolate from the pre-1900 trends. Such a casual impression can be "read into" the graphs in Figures 1-5, though in most instances it is equally easy to confirm an impression of the decrement accruing during 1900 lasting through 1914. However, both impressions are uncritical. Why must one assume that a somewhat linear trend from 1881 to 1899 should continue from 1900 to 1913 or 1914? The answer is of course that one need not. In fact, to do so is a matter of faith. One could argue that the new Civil Code was instrumental in preventing an exponential increase of the divorce and petition for reconciliation rates after 1900.* But to argue either point goes beyond the data. Without comparable "control groups" -- states like those in the German Empire whose divorce laws were not revised in 1900 -- no unequivocal answer can be given to the question "What would the post-1900 trend of the divorce and petition for reconciliation rates have been?"

Previous discussions of the data in Scheidung und Scheidungsrecht by Rheinstein and Wolf, Lüke and Hax have discredited the one conclusion which can be drawn with defensible validity. The time-series quasi-experiment rivals the completely randomized experimental design for validity in some instances. But the inference which enjoys a healthy measure of validity concerns an instantaneous shift in the level of the time-series at the introduction of the experimental treatment and not suppositions about how the time-series should behave long after the treatment has been introduced. There appears to be little doubt that the revision of the divorce laws in 1900 did produce statistically significant effects on both the divorce and petition for reconciliation rates. Whether the effects were temporary or relatively permanent cannot be determined validly from the available data. The supposition that the effects were temporary should not be cited as though it somehow calls into question the one conclusion for which convincing evidence exists, namely that both the divorce and petition for reconciliation rates show the evidence of adoption of the new Civil Code in 1900.

---

*In the U.S. the divorce rate was rising at a faster rate during the second half of the period from 1887-1917 than it was during the first half of that period. The divorce rate rose 31 points (47 to 78) from 1887 to 1902, but it rose 42 points (78 to 120) from 1902 to 1917. A comparison of the pre-1900 and post-1900 trend lines for both the divorce rate and the petition for reconciliation rate in Germany reveals about the same acceleration of the rates after 1900.

References

Box, G. E. P. and Tiao, George C. A change in level of a non-stationary time series. Biometrika, 1965, 52, 181-192.

Campbell, Donald T. From description to experimentation: interpreting trends as quasi-experiments. Chapter 12 in Harris, Chester W. (Ed.), Problems in Measuring Change. Madison, Wisconsin: University of Wisconsin Press, 1963.

Campbell, Donald T. and Stanley, Julian C. "Experimental and quasi-experimental designs for research on teaching." Chapter 5 in N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Maguire, Thomas O. and Glass, Gene V. A program for the analysis of certain time-series quasi-experiments. Educational and Psychological Measurement, 1967, 27, 743-750.

Rheinstein, Max. Divorce and the law in Germany: a review. American Journal of Sociology, 1959, 65, 489-498.

Willcox, Walter E. The Divorce Problem: A Study in Statistics. New York. (Second Edition), 1897.

Wolf, E., Lüke, G., and Hax, H. Scheidung und Scheidungsrecht: Grundfrägen der Ehescheidung in Deutschland. Tübigen: J. C. B. Mohr, 1959.

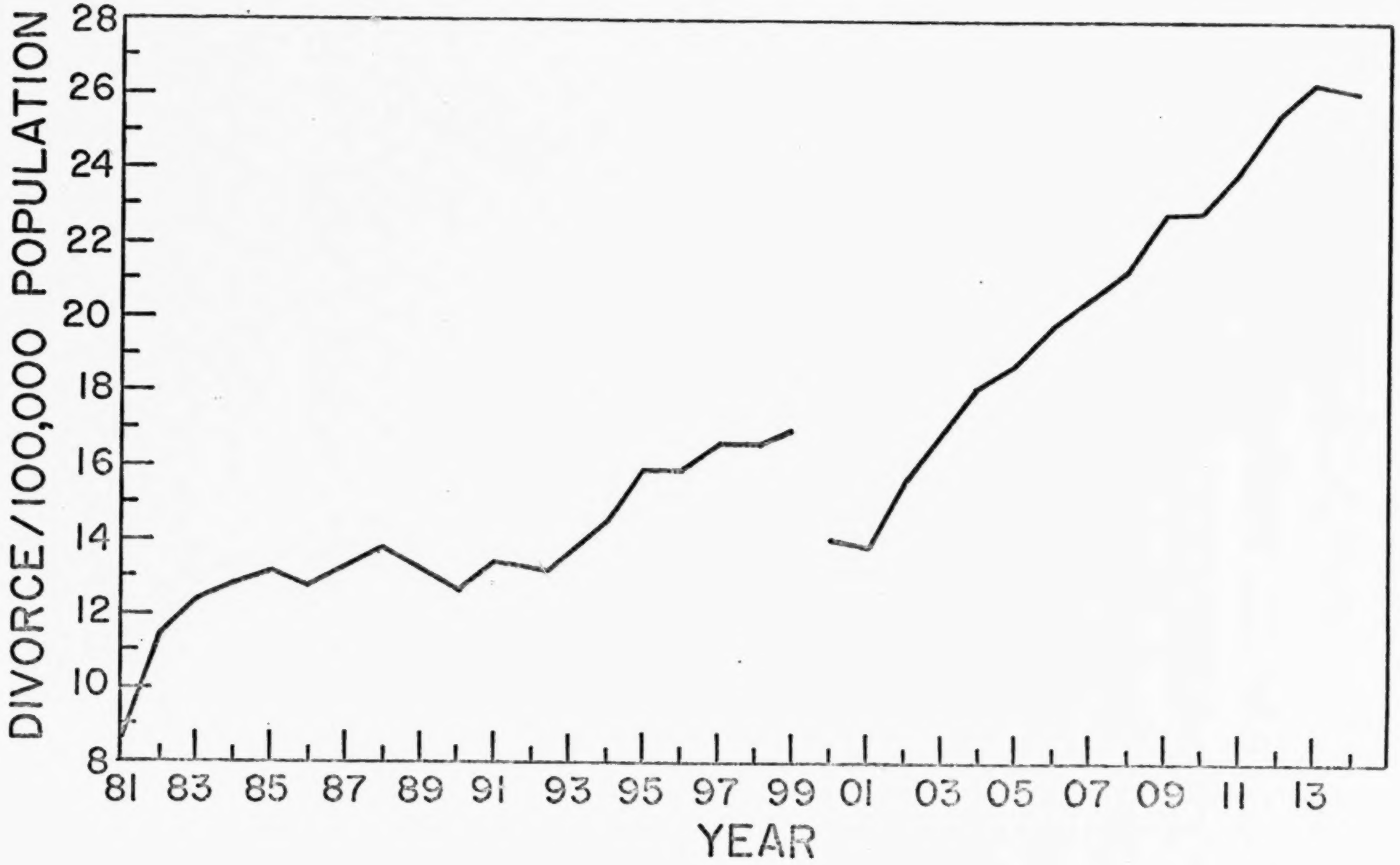


Figure I. DIVORCE RATE FOR GERMAN EMPIRE (1881-1914).

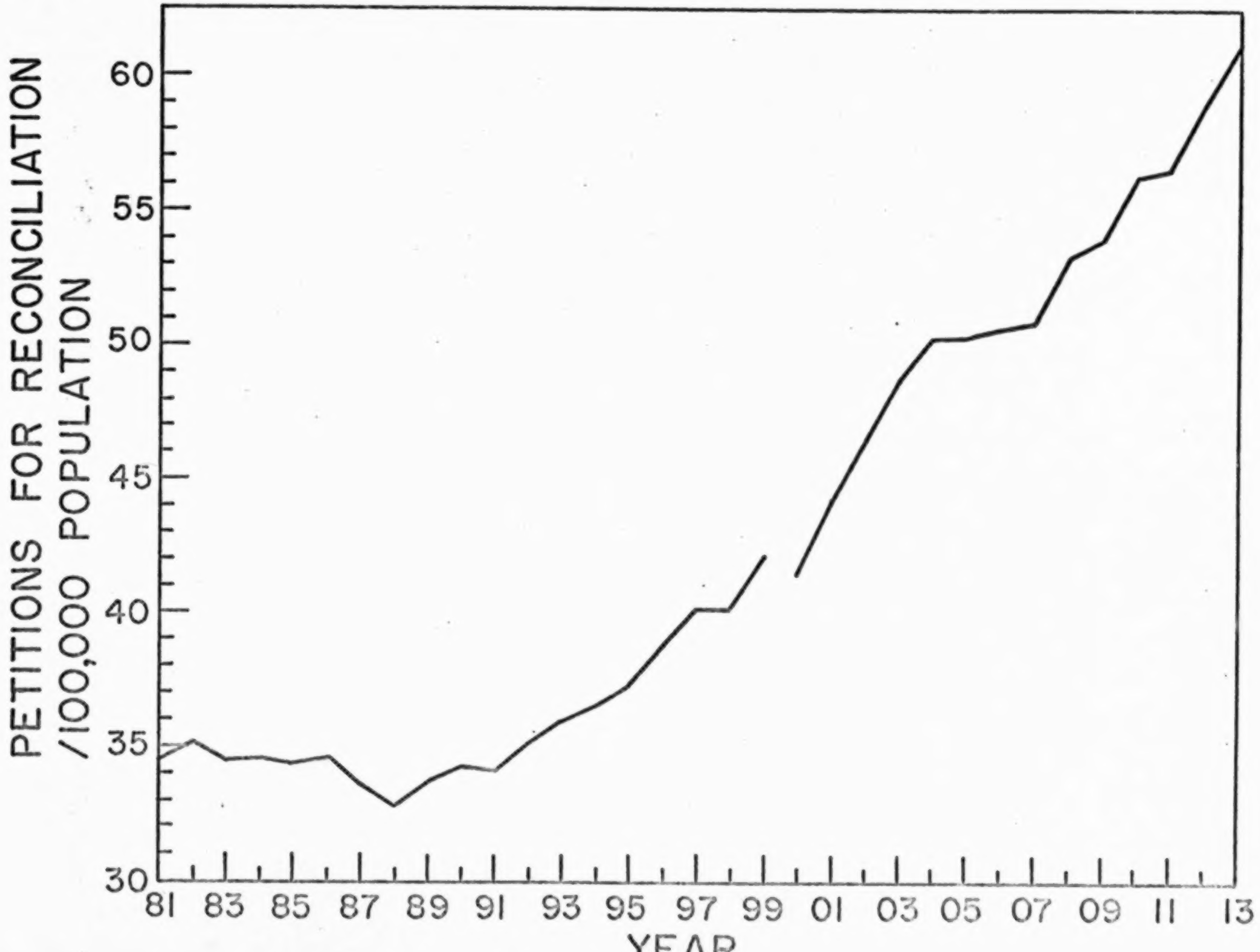


Figure 2. PETITION FOR RECONCILIATION RATE FOR GERMAN EMPIRE (1881-1913)

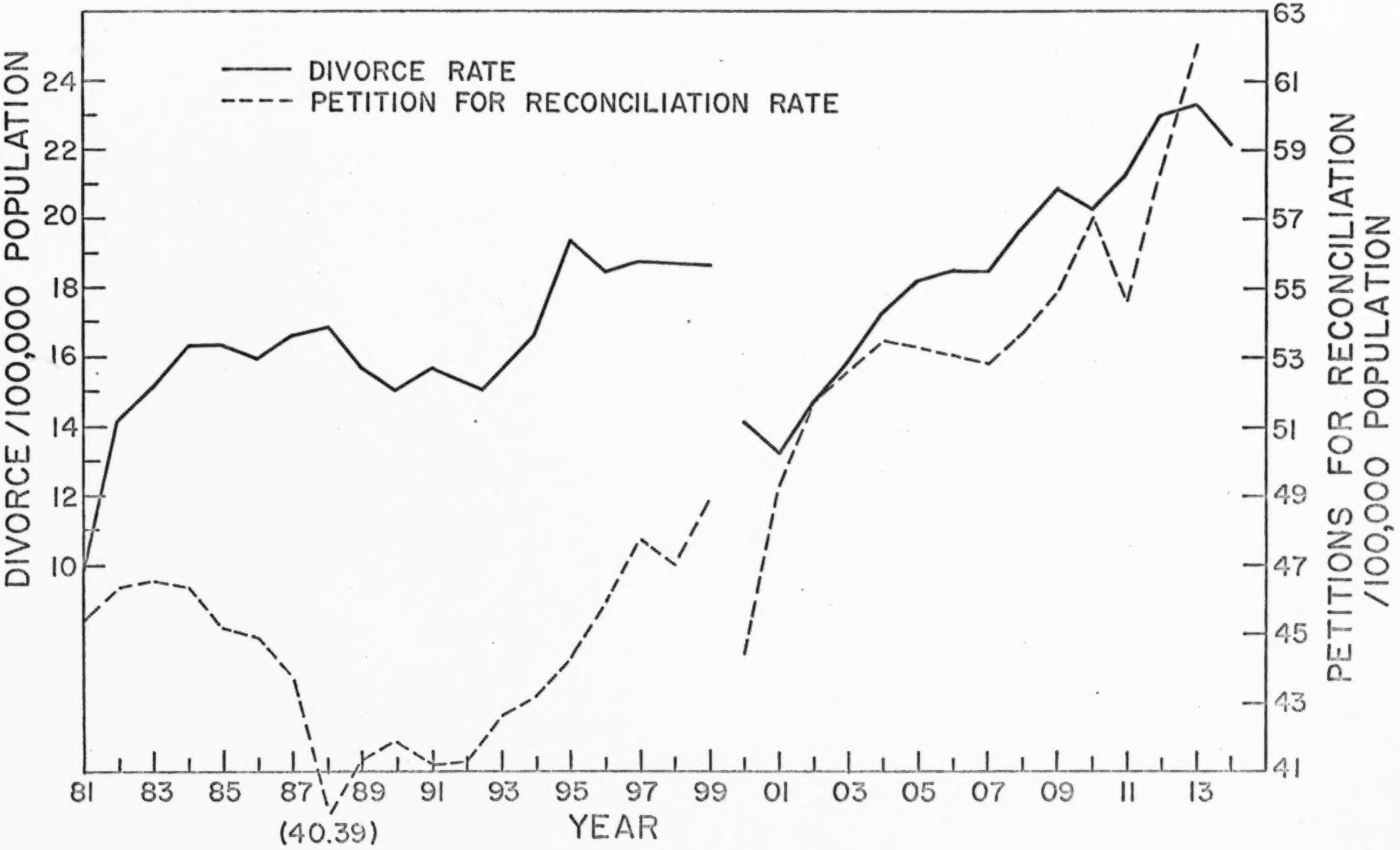


Figure 3 DIVORCE AND PETITION FOR RECONCILIATION RATES FOR STATES UNDER PRUSSIAN CODE PRIOR TO 1900

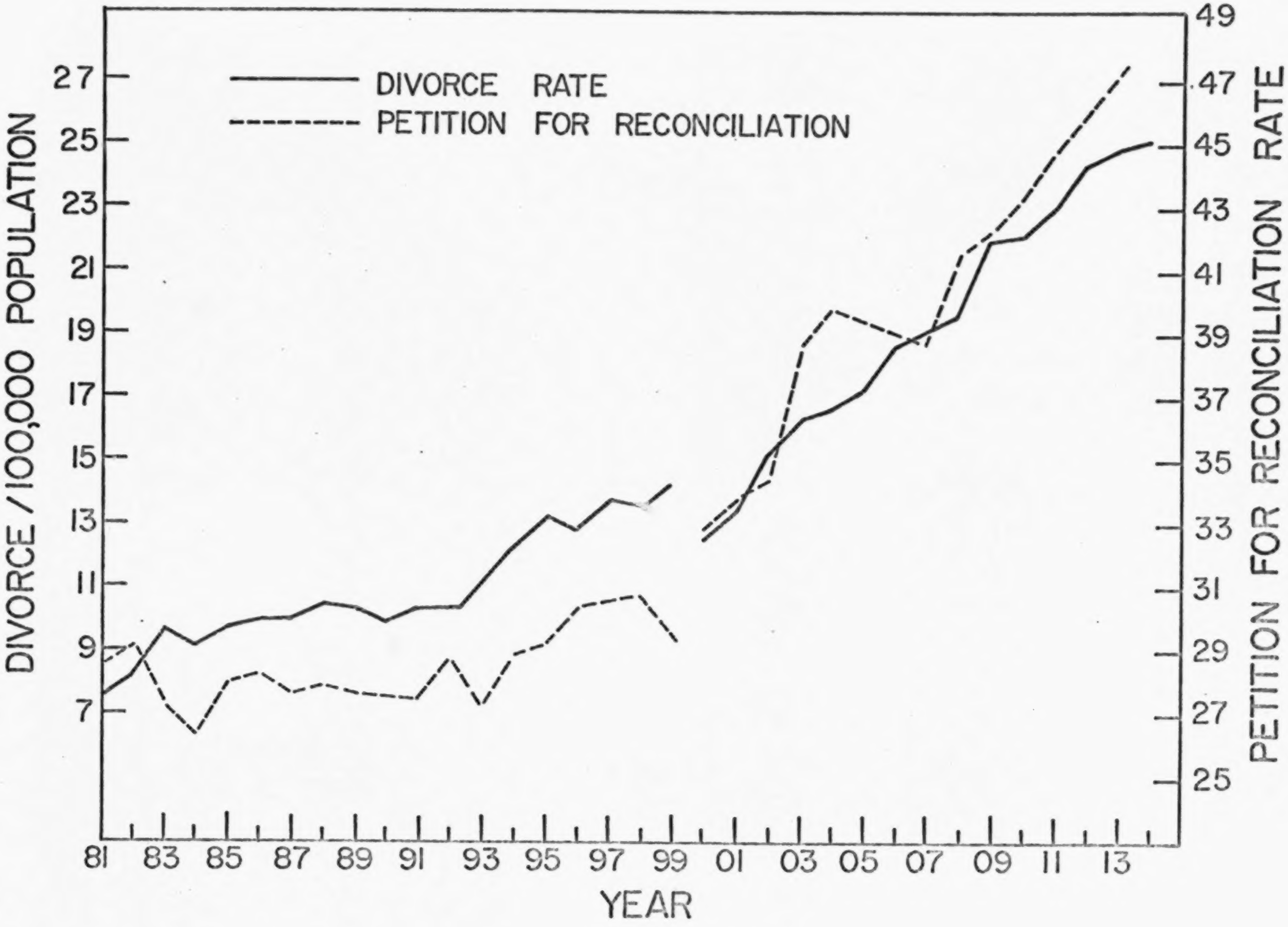


Figure 4. DIVORCE AND PETITION FOR RECONCILIATION RATES FOR STATES UNDER COMMON LAW PRIOR TO 1900

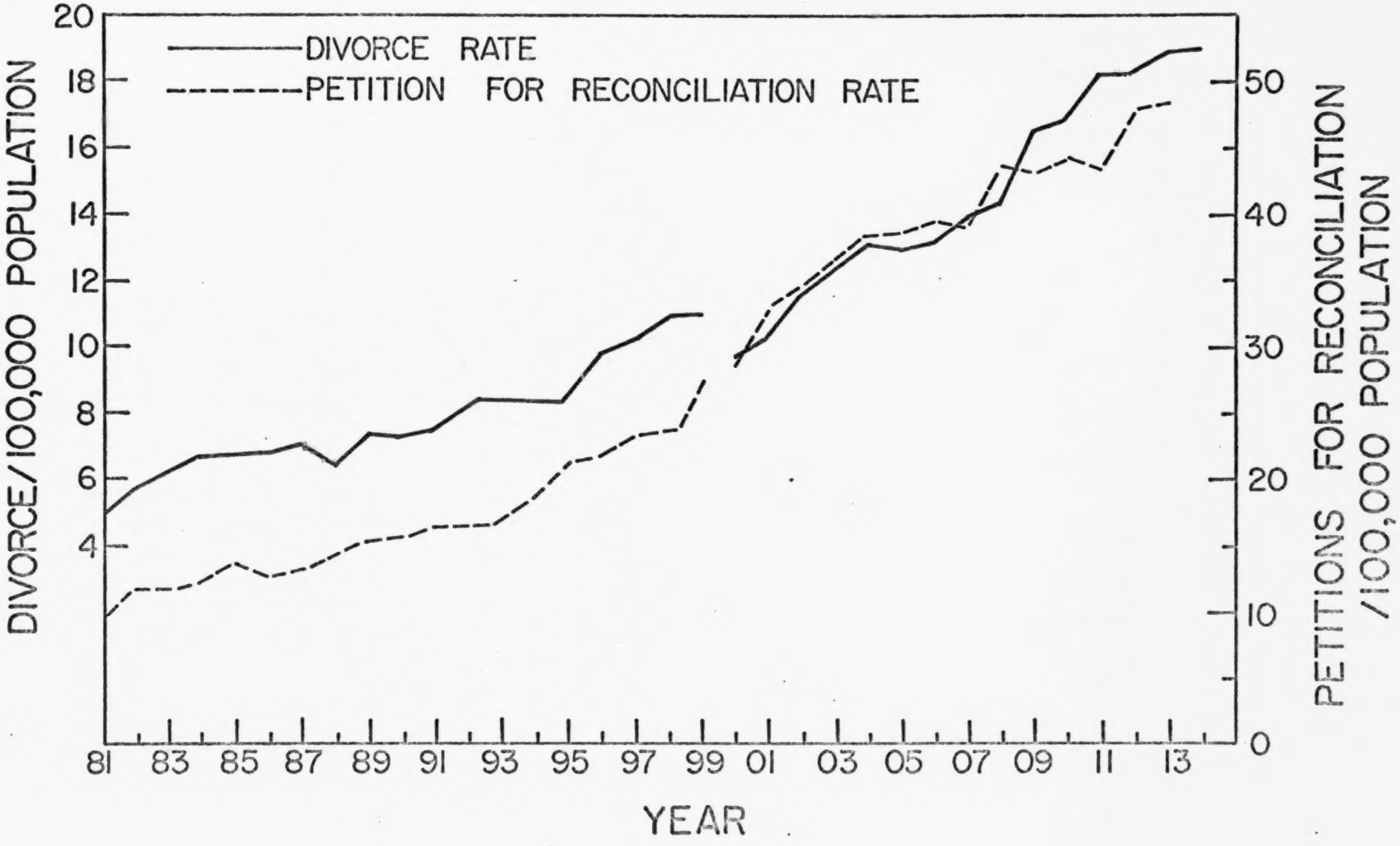


Figure 5. DIVORCE AND PETITION FOR RECONCILIATION RATES FOR STATES UNDER CODE NAPOLEON PRIOR TO 190

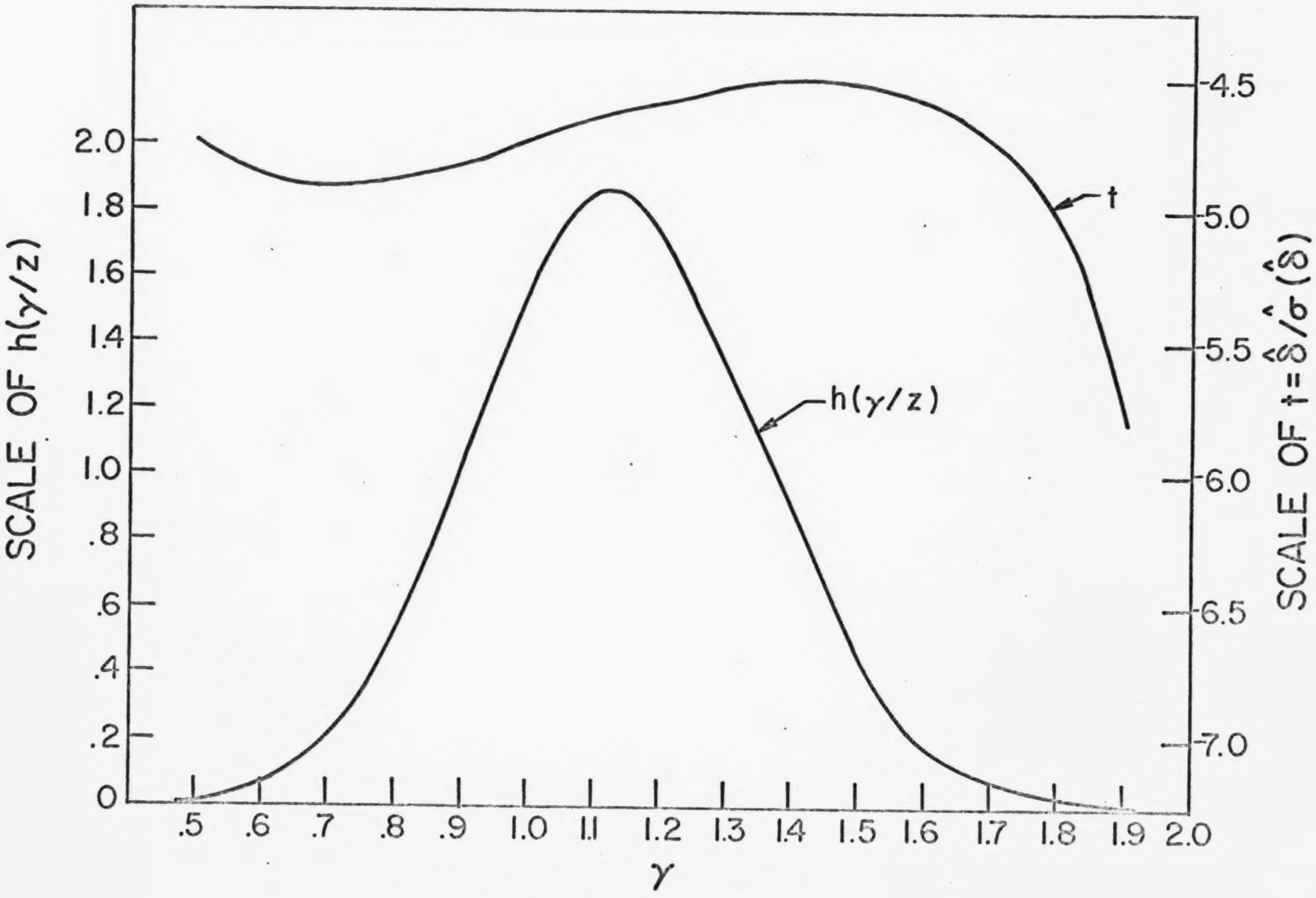


Figure 6. ANALYSIS FOR CHANGE IN LEVEL AT 1900 OF DIVORCE RATE FOR GERMAN EMPIRE (DATA IN FIGURE 1)

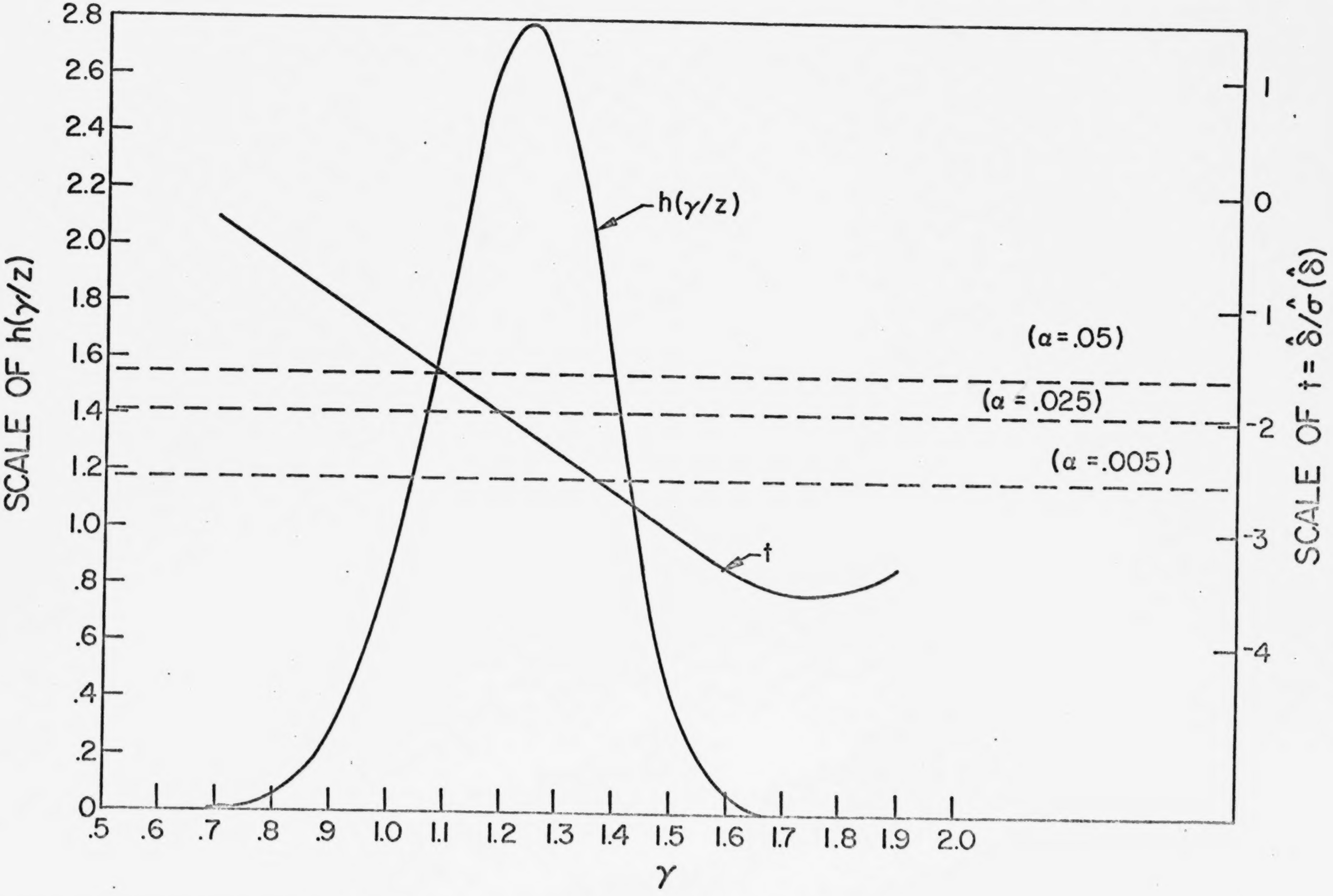


Figure 7. ANALYSIS FOR CHANGE IN LEVEL AT 1900 OF PETITION FOR RECONCILIATION RATE FOR GERMAN EMPIRE (DATA IN FIGURE 2).

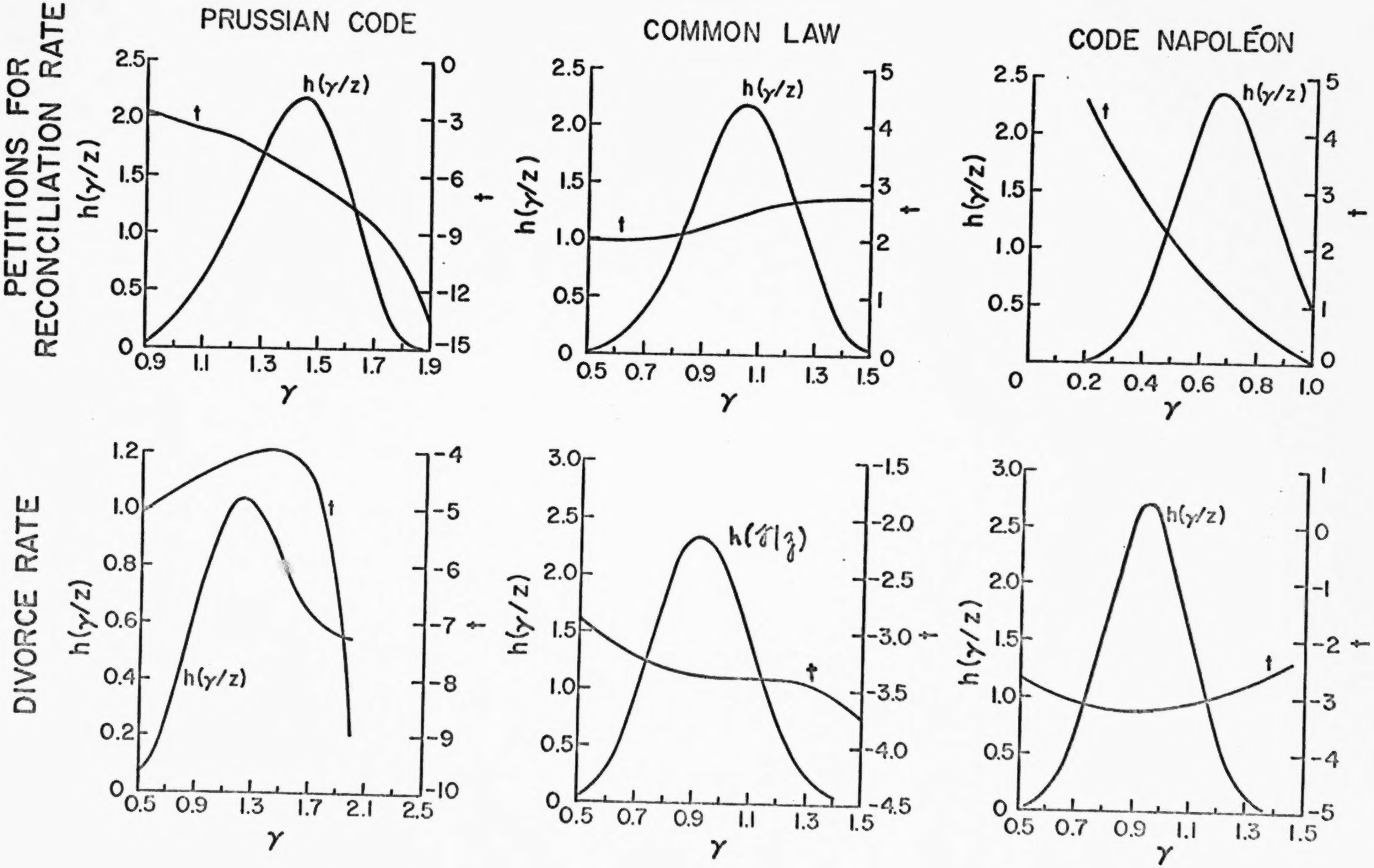


FIGURE 8. ANALYSES FOR CHANGE IN LEVEL AT 1900 OF DIVORCE AND PETITION FOR RECONCILIATION RATES FOR STATES UNDER THREE DIFFERENT LEGAL SYSTEMS PRIOR TO 1900 (DATA IN FIGURES 3–5).

Chapter V

AN ANALYTIC TECHNIQUE EMPLOYING A CHANGE IN DRIFT OF A TIME-SERIES ASSOCIATED WITH THE INTRODUCTION OF A TREATMENT

The model of Chapter IV with its procedures for the estimation and significance testing of both the constant drift, μ, of a time-series and its instantaneous shift in level, δ, at the introduction of a treatment provides a powerful tool for analyzing a large class of time-series experiments. In this chapter, we shall report on an even more general model and analysis which were developed by Dr. George C. Tiao for use in this project.

The Problem

It may occur that the effect of the introduction of a treatment, _T_, into a time-series does not result in an instantaneous change in the level of the series but does change the direction of its drift. The series of observations in Figure 1 evidence no change in level at _T_ but a change in direction of drift. In Figure 2, both changes at _T_ are present.

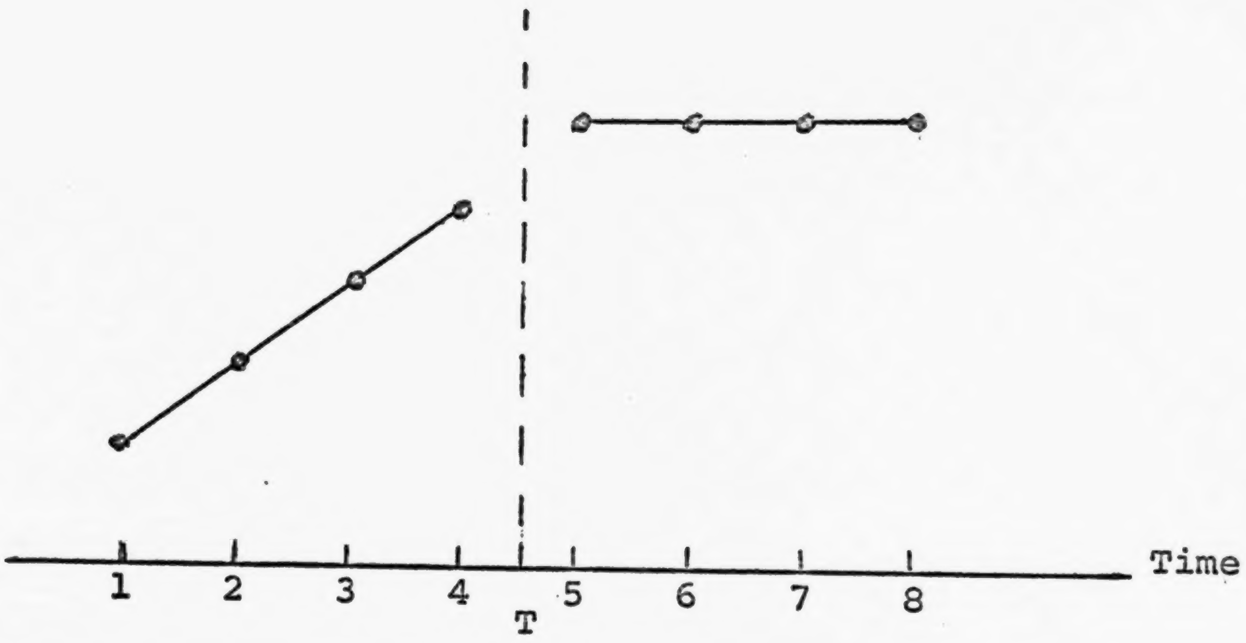

Time
1
2
3
4
T
5
6
7
8

Figure 1. A Time-Series Showing a Change of Direction of Drift but No Change in Level at T.

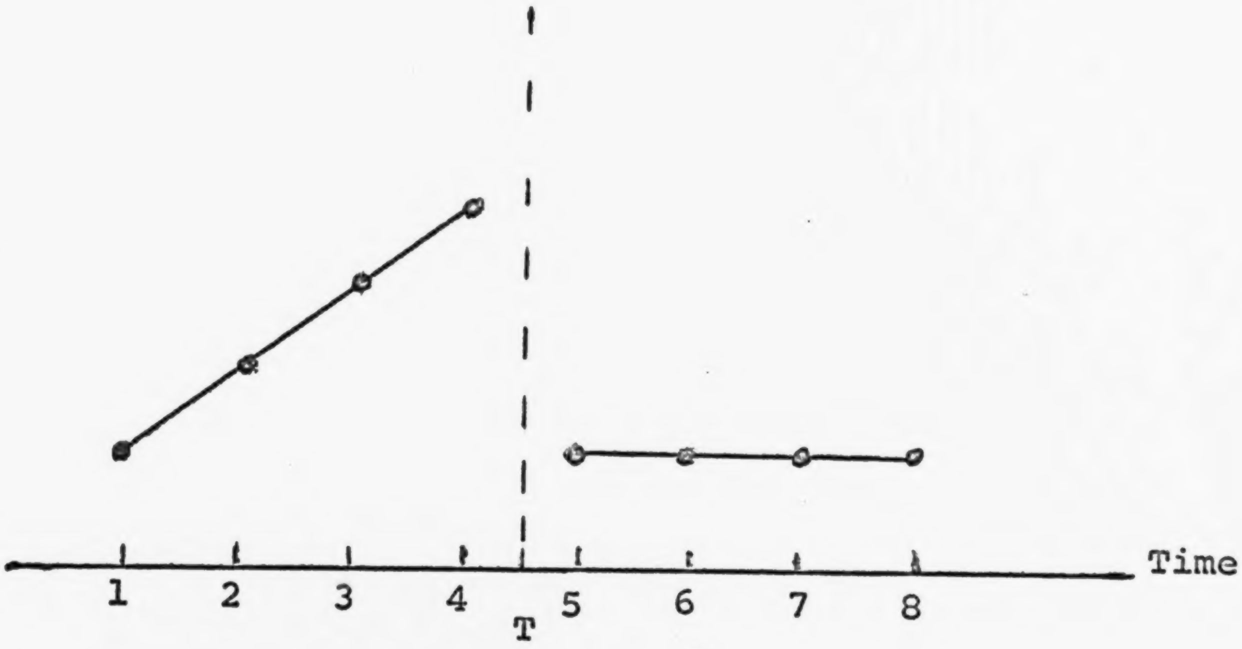

Time
1
2
3
4
T
5
6
7
8

Figure 2. A Time-Series Showing Both a Change in Level and a Change of Direction of Drift at T.

The model of Chapter IV can be modified so that a parameter descriptive of a change in μ, the drift of the series, is incorporated. We shall see how it is then possible to estimate all of the parameters in the model for a given value of γ and to test hypotheses about each. Finally, the likelihood distribution of γ can be found for a set of n observations for use in inferential analyses of δ, μ, and the change in μ.

The Model

Let z_t denote the observation of a series at time t. The following model is proposed for the n_1 observations prior to the introduction of a treatment T:

$$z_t = L + \gamma\mu(t-1) + \mu + \gamma + \gamma \sum_{j=1}^{t-1} \alpha_j + \alpha_t, \qquad (1)$$

where the interpretation of the elements of the model are identical to their interpretation in Chapter IV. The following model is put forward as descriptive of the behavior of the series for the n_2 observations following the introduction of T:

$$z_t = L + \gamma\mu(t-1) + \mu + \gamma\Delta(t-n_1-1) + \Delta + \gamma \sum_{j=1}^{t-1} \alpha_j + \alpha_t + \delta, \qquad (2)$$

where δ is the change of level of the series between times $\underline{n}_1$ and $\underline{n}_1 + 1$, and Δ is the change in the drift of the series between these two times. Prior to $\underline{T}$, the series drifts (on the average) at a rate of $\gamma\mu$ units (up or down depending on the sign of μ) for each unit of time; after $\underline{T}$, the series drifts $\gamma(\mu + \Delta)$ units on the average for each unit of time.

Analytic Procedures

As before, a collection of $\underline{n}_1 + \underline{n}_2$ observations of z are made; these values of z_t are then transformed for a given value of γ as follows:

$$y_1 = z_1$$

$$y_t = z_t - \gamma \sum_{j=1}^{t-1} (1-\gamma)^{j-1} z_{t-j} \qquad \text{for } \underline{t} = 2, \ldots, n_1 + n_2. \tag{3}$$

The $(\underline{n}_1 + \underline{n}_2)$ by 1 vector of y's can be expressed as a linear model in terms of the design matrix $\underline{X}$, the vector of parameters $\Theta^T = (L, \delta, \mu, \Delta)$ and the vector, e, of observations of a random normal variable with variance σ^2, as in Chapter IV:

$$y = X\Theta + e$$

$$\left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_{n_1-1} \\ y_{n_1} \\ \hdashline y_{n_1+1} \\ \vdots \\ y_{n_1+n_2-1} \\ y_{n_1+n_2} \end{array}\right] = \left[\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 1 & 0 & (1-\gamma) & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & (1-\gamma)^{n_1-2} & 0 \\ 1 & 0 & (1-\gamma)^{n_1-1} & 0 \\ \hdashline 1 & 1 & (1-\gamma)^{n_1} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & (1-\gamma)^{n_1+n_2-2} & (1-\gamma)^{n_2-2} \\ 1 & 1 & (1-\gamma)^{n_1+n_2-1} & (1-\gamma)^{n_2-1} \end{array}\right] \begin{bmatrix} \mu \\ \Delta \\ L \\ \delta \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{n_1+n_2} \end{bmatrix}$$

For a single value of γ, the least-squares estimates of the parameters in Θ are obtained from the equation

$$\hat{\Theta} = \begin{bmatrix} \hat{\mu} \\ \hat{\Delta} \\ \hat{L} \\ \hat{\delta} \end{bmatrix} = (X^TX)^{-1} X^Ty. \tag{4}$$

The "residual variance" in fitting the model in (1) and (2) to the observations z_t is given by

$$s^2 = [(y - X\hat{\Theta})^T(y - X\hat{\Theta})]/(n_1 + n_2 - 4). \tag{5}$$

The following distributional statements about the estimates of the parameters follow from the assumption of normality of α_t and traditional sampling theory:

$$\frac{\hat{\mu} - \mu}{s\sqrt{c^{11}}} \sim t_{n_1+n_2-4},$$

$$\frac{\hat{\Delta} - \Delta}{s\sqrt{c^{22}}} \sim t_{n_1+n_2-4}, \tag{6}$$

$$\frac{\hat{L} - L}{s\sqrt{c^{33}}} \sim t_{n_1+n_2-4},$$

$$\frac{\hat{\delta} - \delta}{s\sqrt{c^{44}}} \sim t_{n_1+n_2-4}, \text{ where}$$

c^{jj} is the $\underline{j}$th diagonal element of $(X^TX)^{-1}$.

The above calculations are performed for a single value of γ which is restricted to the open interval (0,2). Since γ is generally unknown, information regarding its likely values must be found from the data themselves. The likelihood distribution of γ given the $n_1 + n_2$ observations z_t is given -- to within a constant of proportionality -- by formula (7):

$$h(\gamma|z) \propto |X^T X|^{-1/2} s^{-(n_1+n_2-3)}. \tag{7}$$

Suppose one wishes to test the following hypotheses at the α-level of significance:

H_0: $\delta = 0$, H_0: $\Delta = 0$.

H_1: $\delta \neq 0$, H_1: $\Delta \neq 0$.

In this instance, the investigator wishes to ascertain whether or not the introduction of a treatment between time n_1 and time $n_1 + 1$ results in either an instantaneous shift of level of the series, an instantaneous shift in the direction of the drift of the series, or both.

For each value of γ between 0 and 2, the values of $\hat{\delta}$, $\hat{\Delta}$ and s^2 are obtained using formulas (4) and (5). The null hypotheses H_0: $\delta = 0$ and H_0: $\Delta = 0$ are tested with the t-statistics $t = \hat{\delta}/(s\sqrt{c^{44}})$ and $t = \hat{\Delta}/(s\sqrt{c^{22}})$ which both have Student's t-distribution with $n_1 + n_2 - 4$ degrees of freedom if the null hypotheses are true. These values of t are likewise calculated for each value of γ between 0 and 2.

Next, the likelihood distribution of γ is determined and the graph of $\underline{t}$ for $\hat{\delta}$ and the graph of $\underline{t}$ for $\hat{\Delta}$ are inspected over the range of γ's which are shown to be likely by inspection of $h(\gamma|z)$. It is determined whether the value of $\underline{t}$ for either $\hat{\delta}$ or $\hat{\Delta}$ tends to be significant [i.e., outside the $100(\alpha/2)$ and $100(1 - \alpha/2)$ percentile points in the $\underline{t}$-distribution with $\underline{n}_1 + \underline{n}_2 - 4$ df] or nonsignificant over the range of likely values of γ. Of course, very large (in absolute value) values of $t = \hat{\delta}/(s\sqrt{c^{44}})$ or $t = \hat{\Delta}/(s\sqrt{c^{22}})$ lead to rejection of H_0: $\delta = 0$ or H_0: $\Delta = 0$, respectively.

Computer Programs

Computer programs for the analyses in Chapter IV and this chapter appear as Appendices A and B, respectively, to this report. Source decks for both programs are available upon request from the authors.

Appendix A

Computer Program for Analysis of Time-Series Experiment with Constant Drift

Appendix B

Computer Program for Analysis of Time-Series Experiment with Possible Change in Drift

```
C TIME SERIES INCORPORATING CHANGE IN DRIFT
C PARAMETER CARD ONE  COLS. 1,2  NB=NUMBER OF PROBLEMS TO BE RUN
C CARDS 2,3,4 ARE REQUIRED FOR EACH PROBLEM TO BE RUN
C CARD 2 COLS.2-80  TITLE OF PROBLEM
C CARD 3 COLS 1-80 FORMAT FOR DATA CARDS..IN THE FORM (FORMAT)
C CARD 4 COLS. 1-4 NUMBER OF PRE TREATMENT MEASURES. COLS. 5-8 NUMBER OF
C POST TREATMENT MEASURES.  COL. 9, 1 IF CORRELOGRAM REQUIRED, 0 OTHERWISE
C CARD 5 ETC. DATA CARDS FOR FIRST PROBLEM NO SEPERATION BETWEEN PRE AND POST
C TREATMENT DATA
      DIMENSION Z(500),Y(50( ),X(500,4),XTXIN(4,4),XTX(4,4),XTY(4)
      DIMENSION THETA(4),FMT(18),TITLE(18),PD(200),XOUT(201,11)
      DIMENSION SE(4),T(4),B(4,1)
      DIMENSION XX(200,10),DELCON(200,5),FNAM(5,5),G1(5),G2(5),G3(5),G4(
     15),G5(5),G6(5),G7(5),G8(5),G9(5),G10(5)
      DATA G1/20HT..CHANGE IN LEVEL  /
      DATA G2/20HT..CHANGE IN SLOPE  /
      DATA G3/20HSCALED POSTERIOR    /
      DATA G4/20HLOWER 99 PERCENT    /
      DATA G5/20HLOWER 95 PERCENT    /
      DATA G6/20HDELTA               /
      DATA G7/20HUPPER 95 PERCENT    /
      DATA G8/20HUPPER 99 PERCENT    /
      XXCK=1.0E-15
      NBB=0
      READ(5,36)NB
  36  FORMAT(I2)
  39  READ(5,1)(TITLE(I),I=1,18)
   1  FORMAT(18A4)
      JK=0
      NIT=0
      G=0.01
      NBB=NBB+1
      WRITE(6,37)NBB
  37  FORMAT(//9H PROBLEM I2//)
      WRITE(6,1)(TITLE(I),I=1,18)
      READ(5,1)(FMT(I),I=1,18)
      READ(5,3)N1,N2,NCC
   3  FORMAT(2I4,I1)
      NTOT=N1+N2
      READ(5,FMT)(Z(I),I=1,NTOT)
      WRITE(6,601)N1,N2
  601 FORMAT(//8X,4H N1=I3,4H N2=I3//)
      WRITE(6,602)
 602  FORMAT(15H     INPUT DATA)
      WRITE(6,600)(Z(I),I=1,N1)
 600  FORMAT(5(1XE13.5))
      NNN=N1+1
      WRITE(6,603)(Z(I),I=NNN,NTOT)
 603  FORMAT(//5(1XE13.5))
      IF(NCC)702,701,702
  702 CALL CORREL (N1,N2,Z)
  701 CONTINUE
      WRITE(6,666)
  666 FORMAT(1H0)
      WRITE(6,22)
  22  FORMAT(1X114H        RESIDUAL               T FOR     CHANGE IN  T
     1 FOR                  T FOR     CHANGE IN    T FOR     SCALED)
      WRITE(6,2022)
 2022 FORMAT(1X117HGAMMA   VARIANCE    LEVEL      LEVEL       LEVEL
     1CHANGE      SLOPE     SLOPE       SLOPE      CHANGE    POSTERIOR)
   25 CONTINUE
C CALCULATION OF  Y SCORES FROM THE DATA ..CHECK FOR  UNDERFLOW WHEN GAMMA
C IS 1.0 AND  Y(N) IS NEARLY EQUAL TO Y(N-1)
      Y(1)=Z(1)
      DO 5 I=2,NTOT
      II=I-1
      YY=ABS(Y(II))
      IF(YY-XXCK)42,42,45
   45 IF(YY-.000001)40,40,41
   40 GG=ABS(1.0-G)
      IF(GG-.001)42,42,41
   42 Y(I)=Z(I)-Z(II)
      GO TO 5
   41 Y(I)=(Z(I)-Z(II))+(1.0-G)*Y(II)
    5 CONTINUE
C CALCULATION OF WEIGHTS, IF ABSOLUTE VALUE OF X IS LESS THAN 1.0E-15, THEN I
C IS SET EQUAL TO ZERO TO PREVENT UNDERFLOW
      DO 1000 I=1,NTOT
 1000 X(I,1)=1.0
      DO 2000 I=1,N1
 2000 X(I,2)=0.0
      NNN1=N1+1
      DO 2001 I=NNN1,NTOT
 2001 X(I,2)=1.0
      X(1,3)=1.0
      X(2,3)=1.0-G
      DO 6 I=3,NTOT
      II=I-1
      X(I,3)=X(2,3)*X(II,3)
      XXX=ABS(X(I,3))
      IF(XXCK-XXX)6,6,32
   32 X(I,3)=0.0
    6 CONTINUE
      DO 7 I=1,N1
    7 X(I,4)=0.0
      NN=N1+1
      DO 8 I=NN,NTOT
      II=I-N1
      X(I,4)=X(II,3)
      XXX=ABS(X(I,4))
      IF(XXCK-XXX)8,8,33
   33 X(I,4)=0.0
    8 CONTINUE
      DO 11 I=1,4
      DO 11 J=1,4
      XTX(I,J)=0.0
   11 XTXIN(I,J)=0.0
C CALCULATION OF X TRANSPOSE X INVERSE
      DO 4050 I=1,4
      DO 4050 J=1,4
      DO 4050 K=1,NTOT
 4050 XTX(I,J)=XTX(I,J)+X(K,I)*X(K,J)
      DO 2002 I=1,4
      DO 2002 J=1,4
 2002 XTXIN(I,J)=XTX(I,J)
      DO 4061 I=1,4
 4061 XTY(I)=0.0
      DO 4062 I=1,4
      DO 4062 J=1,NTOT
 4062 XTY(I)=XTY(I)+X(J,I)*Y(J)
```

```
      DO 2003 I=1,4
 2003 B(I,1)=XTY(I)
      CALL MATINV(XTXIN,4,B,1,DET)
      DO 2020 I=1,4
 2020 THETA(I)=B(I,1)
      DO 2004 I=1,4
 2004 SE(I)=XTXIN(I,I)
      FNTOT=NTOT
      YTY=0.0
C CALCULATION OF THE RESIDUAL VARIANCE
      DO 18 I=1,NTOT
  18  YTY=YTY+Y(I)**2
      DO 31 I=1,4
  31  XTY(I)=0.0
      DO 19 J=1,4
      DO 19 I=1,4
  19  XTY(J)=XTY(J)+THETA(I)*XTX(I,J)
      FITVAR=0.0
      DO 20 I=1,4
  20  FITVAR=FITVAR+XTY(I)*THETA(I)
      S=YTY-FITVAR
      S=S/(FNTOT-4.0)
C CALCULATION OF THE STANDARD ERRORS OF DELTA AND MU
      DO 2010 I=1,4
      SE(I)=SQRT(S*SE(I))
 2010 T(I)=THETA(I)/SE(I)
C CALCULATION OF THE POSTERIOR DISTRIBUTION... LOGS ARE USED TO PREVENT
C OVERFLOW
      SK=ALOG(S)
      DET=ALOG(DET)
      H=(-.5*DET)-(.5*(FNTOT-4.0)*SK)
      H=.4342945*H
      JK=JK+1
      XOUT(JK,1)=G
      XOUT(JK,2)=S
      XOUT(JK,3)=THETA(3)
      XOUT(JK,4)=T(3)
      XOUT(JK,5)=THETA(4)
      XOUT(JK,6)=T(4)
      XOUT(JK,7)=THETA(1)
      XOUT(JK,8)=T(1)
      XOUT(JK,9)=THETA(2)
      XOUT(JK,10)=T(2)
      IF(NTOT-30)1004,1005,1005
 1005 DELCON(JK,1)=THETA(4)-2.58*SE(4)
      DELCON(JK,2)=THETA(4)-1.96*SE(4)
      DELCON(JK,3)=THETA(4)
      DELCON(JK,4)=THETA(4)+1.96*SE(4)
      DELCON(JK,5)=THETA(4)+2.58*SE(4)
 1004 NIT=NIT+1
      PD(NIT)=H
C INCREMENT GAMMA BY .01 AND ITERATE
      G=G+1.00000000E-02
      IF(NIT-199)30,30,26
  30  GO TO 25
  26  CONTINUE
C FIND MAXIMUM VALUE OF THE POSTERIOR
      FIN=PD(1)
      DO 506 I=2,199
      IF(FIN-PD(I))505,506,506
 505  FIN=PD(I)
 506  CONTINUE
C RESCALE POSTERIOR BY DIVIDING ALL VALUES OF THE POSTERIOR BY THE MAX VALUE
      DO 507 I=1,199
      PD(I)=PD(I)-FIN
      YY=ABS(PD(I))
      IF(YY-35.0)509,508,508
 508  PD(I)=0.0
      GO TO 507
 509  PD(I)=PD(I)/.4342945
      PD(I)=EXP(PD(I))
 507  CONTINUE
C CONVERT AREA OF THE POSTERIOR DISTRIBUTION TO UNIT AREA BY METHOD OF TRAPEZ
      AREA=0.0
      DO 511 I=2,199
      II=I-1
  511 AREA=AREA+.005*(PD(I)+PD(II))
      DO 512 I=1,199
  512 PD(I)=PD(I)/AREA
      DO 513 I=1,199
  513 XOUT(I,11)=PD(I)
      DO 514 I=1,199
  514 WRITE(6,23)(XOUT(I,J),J=1,11)
  23  FORMAT(1X,1F5.2,10(1XE10.3))
      DO 700 I=1,199
      XX(I,1)=XOUT(I,5)
      XX(I,2)=XOUT(I,9)
 700  XX(I,3)=XOUT(I,11)
      M=199
      N=3
      ISCALE=1
C TITLE FOR  PLOT OF THE POSTERIOR DISTRIBUTION AND FOR STUDENT T
      DO 760 I=1,5
      FNAM(1,I)=G1(I)
      FNAM(2,I)=G2(I)
 760  FNAM(3,I)=G3(I)
      CALL PLOT(M,N,ISCALE,XX,FNAM)
      N=5
C PLOT CONFIDENCE INTERVALS AROUND DELTA
      IF(NTOT-30)1007,1006,1006
 1006 DO 1008 I=1,199
      DO 1008 J=1,5
 1008 XX(I,J)=DELCON(I,J)
      DO 756 I=1,5
      FNAM(1,I)=G4(I)
      FNAM(2,I)=G5(I)
      FNAM(3,I)=G6(I)
      FNAM(4,I)=G7(I)
 756  FNAM(5,I)=G8(I)
      WRITE(6,1016)
 1016 FORMAT(//34H CONFIDENCE INTERVALS AROUND DELTA)
      WRITE(6,1017)
 1017 FORMAT(/3X,5HGAMMA,2X,52H  LOWER 99   LOWER 95   DELTA    UPPER 95
     1   UPPER 99)
      DO 1019 I=1,199
 1019 WRITE(6,1018)XOUT(I,1),(DELCON(I,J),J=1,5)
 1018 FORMAT(3X,1F5.2,1X5E10.3)
      WRITE(6,1015)
 1015 FORMAT(//1X,47H GRAPH OF CONFIDENCE INTERVALS AROUND DELTA HAT)
      CALL PLOT(M,N,ISCALE,XX,FNAM)
 1007 NB=NB-1
      IF(NB)38,38,39
   38 CONTINUE
      STOP
      END
      SUBROUTINE CORREL (N1,N2,Z)
      DIMENSION Z(500),RLAG(400),Y(500),XX(200,10),FNAM(5,5)
      DIMENSION G9(5),G10(5)
      DATA G9/20HPRE TREATMENT DATA  /
      DATA G10/20HPOST TREATMENT DATA /
      NTOT=N1+N2
C PREPARATION OF TITLE FOR CORRELOGRAM PLOT
      DO 804 JJJ=1,4
      IND=JJJ
      IF(IND-2)805,806,807
  807 IF(IND-4)805,806,806
 805  CONTINUE
      DO 757 I=1,5
 757  FNAM(1,I)=G9(I)
      GO TO 808
 806  CONTINUE
      DO 758 I=1,5
 758  FNAM(1,I)=G10(I)
  808 CONTINUE
      GO TO (809,810,811,812),IND
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR PRE TREATMENT DATA
  809 NLAG=N1
      NLOW=1
      NTOP=N1
      WRITE(6,813)
  813 FORMAT(//38H CORRELOGRAM OF PRE TREATMENT RAW DATA)
      GO TO 814
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR POST TREATMENT DATA
  810 NLAG=N2
      NLOW=N1+1
      NTOP=NTOT
      WRITE(6,815)
  815 FORMAT(//39H CORRELOGRAM OF POST TREATMENT RAW DATA)
  814 II=0
      DO 706 I=NLOW,NTOP
      II=II+1
  706 Y(II)=Z(I)
      GO TO 709
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR PRE TREATMENT
C DIFFERENCES  BETWEEN SUCCESSIVE OBSERVATIONS
  811 NLAG=N1-1
      NLOW=1
      NTOP=N1-1
      WRITE(6,800)
  800 FORMAT(//41H CORRELOGRAM OF PRE TREATMENT DIFFERENCES)
      GO TO 816
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR POST TREATMENT
C DIFFERENCES  BETWEEN SUCCESSIVE OBSERVATIONS
  812 NLAG=N2-1
      NLOW=N1+1
      NTOP=NTOT-1
      WRITE(6,803)
  803 FORMAT(//42H CORRELOGRAM OF POST TREATMENT DIFFERENCES)
  816 II=0
      DO 707 I=NLOW,NTOP
      II=II+1
      IK=I+1
  707 Y(II)=Z(IK)-Z(I)
  709 JJ=NLAG*3/4
      DO 716 K=1,JJ
      NUP=NLAG-K
      CROS=0.0
      SUM1=0.0
      SUM2=0.0
      SUMS1=0.0
      SUMS2=0.0
      DO 703 I=1,NUP
      NIND=I+K
      CROS=CROS+Y(I)*Y(NIND)
      SUM1=SUM1+Y(I)
      SUM2=SUM2+Y(NIND)
      SUMS1=SUMS1+Y(I)*Y(I)
  703 SUMS2=SUMS2+Y(NIND)*Y(NIND)
      FNUP=NUP
      DNUM=CROS-(SUM1*SUM2)/FNUP
      DEN=(SUMS1-(SUM1*SUM1)/FNUP)*(SUMS2-(SUM2*SUM2)/FNUP)
      DEN=SQRT(DEN)
      RLAG(K)=DNUM/DEN
      WRITE(6,704)K,RLAG(K)
  704 FORMAT(1X,5H LAG=I3,3X,3H R=F6.3)
  716 CONTINUE
      ISCALE=1
      N=1
      M=JJ
      DO 650 I=1,JJ
  650  XX(I,1)=RLAG(I)
      CALL PLOT(M,N,ISCALE,XX,FNAM)
  804 CONTINUE
      RETURN
      END
      SUBROUTINE PLOT (M,N,ISCALE,XX,FNAM)
C ADAPTED FROM THE SUBROUTINE-GRAPH- OF -PERSUB- WRITTEN BY J.H.WARD,J-.,
C KATHLEEN DAVIS, AND JANICE BUCHHORN,LACKLAND AIR FORCE BASE,TEXAS
C ADAPTED FOR FORTRAN II BY T.O.MAGUIRE,UNIVERSITY OF ILLINOIS
      DIMENSION RAT(10),FND(5,10),ENCRMT(10),PA(120),FMT(3),PB(10)
      DIMENSION XX(200,10),AMAX(10),AMIN(10),B(10),FNAM(5,5),BB(10)
      DATA BB(1),BB(2),BB(3),BB(4),BB(5),BB(6),BB(7),BB(8),BB(9),BB(10)/
     11H1,1H5,1HD,1H5,1H1,1HO,1HX,1HM,1HH,1HA/
      DATA PCROSS,PBLANK,PDASH,PERIOD/1H*,1H ,1H-,1H./
      DATA FMT(1),FMT(2),FMT(3)/1H ,1H0,1H-/
      IF(N-5)200,201,200
  201 DO 276 I=1,5
  276 B(I)=BB(I)
      GO TO 202
  200 DO 277 I=1,3
      II=I+5
  277 B(I)=BB(II)
  202 CONTINUE
   99 FMS=FMT(ISCALE-1)
   98 DO 100 I=1,N
      AMIN(I)=+1.E+37
  100 AMAX(I)=-1.E+37
C SEARCH FOR MAXIMA AND MINIMA
      ABMIN=AMIN(1)
      ABMAX=AMAX(1)
      DO 101 I=1,M
      DO 101 J=1,N
      COMP=XX(I,J)-AMIN(J)
      IF(COMP)4002,4003,4003
 4002 AMIN(J)=XX(I,J)
 4003 COMP=XX(I,J)-AMAX(J)
      IF(COMP)101,101,4004
 4004 AMAX(J)=XX(I,J)
  101 CONTINUE
      IF(N-5)250,251,250
  251 DO 252 J=1,5
      AMAX(J)=AMAX(5)
  252 AMIN(J)=AMIN(1)
  250 CONTINUE
      N1=N
      DO 108 J=1,N
C COMPUTE RESOLUTION OF GRAPH
  108 RAT(J)=(AMAX(J)-AMIN(J))/110.
      DO 110   J=1,N
      ENCRMT(J)=(AMAX(J)-AMIN(J))/4.
      FND(1,J)=AMIN(J)+.05
      COMP=AMIN(J)
      IF(COMP)4005,4006,4006
 4005 FND(1,J)=FND(1,J)-.10
 4006 FND(5,J)=AMAX(J)-.05
      COMP=AMAX(J)
      IF(COMP)4007,4008,4008
C PREPARE ORDINATE LABELS
 4007 FND(5,J)=FND(5,J)-.10
 4008 FND(2,J)=AMIN(J)+ENCRMT(J)+.05
      COMP=FND(2,J)
      IF(COMP)4009,4010,4010
 4009 FND(2,J)=FND(2,J)-.10
 4010 FND(3,J)=AMIN(J)+(ENCRMT(J)*2.)+.05
      COMP=FND(3,J)
      IF(COMP)4011,4012,4012
 4011 FND(3,J)=FND(3,J)-.10
 4012 FND(4,J)=AMAX(J)-ENCRMT(J)+.05
      COMP=FND(4,J)
      IF(COMP)4013,110,110
 4013 FND(4,J)=FND(4,J)-.10
 110  CONTINUE
C PRINT LEFT HAND LABELS
      WRITE(6,7)((FND(I,J),I=1,5),B(J),J=1,N1)
    7 FORMAT(///1XF6.1,21XF6.1,21XF6.1,21XF6.1,21XF6.1,3XA1)
C PLOT LEFT HAND MARGIN
      DO 4014 I=1,59
 4014 PA(I)=PDASH
      WRITE(6,8)(PA(I),I=1,59)
    8 FORMAT(1X,59A2)
      DO 140 I=1,M
      DO 4015 II=1,120
 4015 PA(II)=PBLANK
      DO 121 IX=27,81,27
  121 PA(IX)=PERIOD
      NCOMP=I
 4016 NCOMP=NCOMP-10
      IF(NCOMP)124,4017,4016
 4017 CONTINUE
      DO 123 IX=6,120,2
 123  PA(IX)=PDASH
C RESCALE DATA POINTS
 124  DO 135 K=1,N
      ZL=(XX(I,K)-AMIN(K))/RAT(K)+1.0
      L=ZL
      IF(L-1)6018,6019,6019
 6018 L=1
 6019 IF(110-L)6020,6021,6021
 6020 L=110
 6021 IF(PA(L).EQ.PBLANK)GO TO 130
      IF(PA(L).EQ.PERIOD)GO TO 130
      IF(PA(L).EQ.PDASH)GO TO 130
      PA(L)=PCROSS
      GO TO 135
  130 PB(K)=B(K)
      PA(L)=PB(K)
  135 CONTINUE
      IF(ISCALE-1)6023,136,6023
    6 FORMAT(A1)
 6023 WRITE(6,6)FMS
C PLOT DATA POINTS
  136 WRITE(6,2)I,PERIOD,(PA(J),J=1,110),PERIOD,I
    2 FORMAT(1X,I3,A1,110A1,A1,I3)
  140 CONTINUE
      DO 6026 I=1,59
 6026 PA(I)=PDASH
      WRITE(6,8)(PA(I),I=1,59)
      WRITE(6,7)((FND(I,J),I=1,5),B(J),J=1,N1)
  142 WRITE(6,3)
    3 FORMAT(1H0,10X,16HPLOT DESCRIPTION/1H0,7X,5HTITLE,10X,9HCHARACTER,
     14X,7HMINIMUM,4X,7HMAXIMUM,4X,10HRESOLUTION)
      DO 122 J=1,N
  122 WRITE(6,4)(FNAM(J,K),K=1,5),B(J),AMIN(J),AMAX(J),RAT(J)
    4 FORMAT(1X,5A4,2X,A1,7XF8.3,3XF8.3,4XF8.3)
      RETURN
      END
      SUBROUTINE MATINV(A,N,B,M,DETERM)
C     MATRIX INVERSION WITH ACCOMPANYING SOLUTION OF LINEAR EQUATIONS
      DIMENSION A(4,4),B(4,1),IPIVOT(4),INDEX(4,2),PIVOT(4)
      EQUIVALENCE (IROW,JROW),(ICOLUM,JCOLUM),(AMAX,T,SWAP)
C     INITIALIZATION
 10   DETERM=1.0
 15   DO 20 J=1,N
 20   IPIVOT(J)=0
 30   DO 550 I=1,N
C     SEARCH FOR PIVOT ELEMENT
 40   AMAX=0.0
 45   DO 105 J=1,N
 50   IF (IPIVOT(J)-1) 60,105,60
 60   DO 100 K=1,N
 70   IF (IPIVOT(K)-1) 80,100,740
 80   IF (ABS (AMAX)-ABS (A(J,K))) 85, 85, 100
 85   IROW=J
 90   ICOLUM=K
 95   AMAX=A(J,K)
 100  CONTINUE
 105  CONTINUE
 110  IPIVOT (ICOLUM)=IPIVOT(ICOLUM)+1
C     INTERCHANGE ROWS TO PUT PIVOT ELEMENT ON DIAGONAL
 130  IF(IROW-ICOLUM) 140,260,140
 140   DETERM=-DETERM
 150   DO 200 L=1,N
 160   SWAP=A(IROW,L)
 170   A(IROW,L)=A(ICOLUM,L)
 200   A(ICOLUM,L)=SWAP
 205   IF(M) 260,260,210
 210   DO 250 L=1,M
 220   SWAP=B(IROW,L)
 230   B(IROW,L)=B(ICOLUM,L)
 250   B(ICOLUM,L)=SWAP
 260   INDEX(I,1)=IROW
 270   INDEX(I,2)=ICOLUM
 310   PIVOT(I)=A(ICOLUM,ICOLUM)
 320   DETERM=DETERM*PIVOT(I)
       IF(PIVOT(I)) 330,720,330
C      DIVIDE PIVOT ROW BY PIVOT ELEMENT
 330   A(ICOLUM,ICOLUM)=1.0
 340   DO 350 L=1,N
 350   A(ICOLUM,L)=A(ICOLUM,L)/PIVOT(I)
 355   IF(M) 380,380,360
 360   DO 370 L=1,M
 370   B(ICOLUM,L)=B(ICOLUM,L)/PIVOT(I)
C      REDUCE NON-PIVOT ROWS
 380   DO 550 L1=1,N
 390   IF(L1-ICOLUM) 400,550,400
 400   T=A(L1,ICOLUM)
 420   A(L1,ICOLUM)=0.0
 430   DO 450 L=1,N
 450   A(L1,L)=A(L1,L)-A(ICOLUM,L)*T
 455   IF(M) 550,550,460
 460   DO 500 L=1,M
 500   B(L1,L)=B(L1,L)-B(ICOLUM,L)*T
 550   CONTINUE
C      INTERCHANGE COLUMNS
 600   DO 710 I=1,N
 610   L=N+1-I
 620   IF(INDEX(L,1)-INDEX(L,2)) 630,710,630
 630   JROW=INDEX(L,1)
 640   JCOLUM=INDEX(L,2)
 650   DO 705 K=1,N
 660   SWAP=A(K,JROW)
 670   A(K,JROW)=A(K,JCOLUM)
 700   A(K,JCOLUM)=SWAP
 705   CONTINUE
 710   CONTINUE
       RETURN
 720   WRITE (6      ,730    )
 730   FORMAT (20H  MATRIX IS SINGULAR)
 740   RETURN
       END
       IF(COMP)4005,4006,4006
```

```
C TIME SERIES WITH CONSTANT DRIFT
C PARAMETER CARD ONE  COLS. 1,2  NB=NUMBER OF PROBLEMS TO BE RUN
C CARDS 2,3,4 ARE REQUIRED FOR EACH PROBLEM TO BE RUN
C CARD 2 COLS.2-80  TITLE OF PROBLEM
C CARD 3 COLS 1-80 FORMAT FOR DATA CARDS..IN THE FORM (FORMAT)
C CARD 4 COLS. 1-4 NUMBER OF PRE TREATMENT MEASURES. COLS. 5-8 NUMBER OF

C CARD 5 ETC. DATA CARDS FOR FIRST PROBLEM NO SEPERATION BETWEEN PRE AND POST
C TREATMENT DATA
      DIMENSION Z(500),Y(500),X(500,3),XTXIN(3,3),XTX(3,3),XTY(3)
      DIMENSION THETA(3),FMT(18),TITLE(18),PD(200),XOUT(201,10)
      DIMENSION XX(200,10),DELCON(200,5),FNAM(5,5)
      DIMENSION G1(5),G2(5),G3(5),G4(5),G6(5),G7(5),G8(5),G9(5),G5(5)
      DATA G1/20HT..CHANGE IN LEVEL  /
      DATA G2/20HSCALED POSTERIOR    /
      DATA G3/20HLOWER 99 PERCENT    /
      DATA G4/20HLOWER 95 PERCENT    /
      DATA G5/20HDELTA               /
      DATA G6/20HUPPER 95 PERCENT    /
      DATA G7/20HUPPER 99 PERCENT    /
      XXCK=1.0E-15
      NBB=0
      READ(5,36)NB
  36  FORMAT(I2)
  39  READ(5,1)(TITLE(I),I=1,18)
   1  FORMAT(18A4)
      JK=0
      NIT=0
      G=0.01
      NBB=NBB+1
      WRITE(6,37)NBB
  37  FORMAT(//9H PROBLEM I2//)
      WRITE(6,1)(TITLE(I),I=1,18)
      READ(5,1)(FMT(I),I=1,18)
      READ(5,3)N1,N2,NCC
   3  FORMAT(2I4,I1)
      NTOT=N1+N2
      READ(5,FMT)(Z(I),I=1,NTOT)
      WRITE(6,601)N1,N2
  601 FORMAT(//8X,4H N1=I3,4H N2=I3//)
      WRITE(6,602)
 602  FORMAT(15H      INPUT DATA)
      WRITE(6,600)(Z(I),I=1,N1)
 600  FORMAT(5(1XE13.5))
      NNN=N1+1
      WRITE(6,603)(Z(I),I=NNN,NTOT)
 603  FORMAT(//5(1XE13.5))
      IF(NCC)702,701,702
  702 CALL CORREL (N1,N2,Z)
  701 CONTINUE
      WRITE(6,666)
  666 FORMAT(1H0)
      WRITE(6,22)
  22  FORMAT(1X,116H   GAMMA  RESIDUAL VAR  L HAT       MU HAT    STD ERR
     1 MU T FOR MU    DELTA  STD ERR DELTA T FOR DELTA POST DIST GAMMA)
  25  CONTINUE
C CALCULATION OF  Y SCORES FROM THE DATA ..CHECK FOR  UNDERFLOW WHEN GAMMA
C IS 1.0 AND  Y(N) IS NEARLY EQUAL TO Y(N-1)
      Y(1)=Z(1)
      DO 5 I=2,NTOT
      II=I-1
      YY=ABS(Y(II))
      IF(YY-XXCK)42,42,45
   45 IF(YY-.000001)40,40,41
   40 GG=ABS(1.0-G)
      IF(GG-.001)42,42,41
   42 Y(I)=Z(I)-Z(II)
      GO TO 5
   41 Y(I)=(Z(I)-Z(II))+(1.0-G)*Y(II)
    5 CONTINUE
C CALCULATION OF WEIGHTS, IF ABSOLUTE VALUE OF X IS LESS THAN 1.0E-15, THEN I
C IS SET EQUAL TO ZERO TO PREVENT UNDERFLOW
      DO 1000 I=1,NTOT
 1000 X(I,1)=1.0
      X(1,2)=1.0
      X(2,2)=1.0-G
      DO 6 I=3,NTOT
      II=I-1
      X(I,2)=X(2,2)*X(II,2)
      XXX=ABS(X(I,2))
      IF(XXCK-XXX)6,6,32
   32 X(I,2)=0.0
    6 CONTINUE
      DO 7 I=1,N1
    7 X(I,3)=0.0
      NN=N1+1
      DO 8 I=NN,NTOT
      II=I-N1
      X(I,3)=X(II,2)
      XXX=ABS(X(I,3))
      IF(XXCK-XXX)8,8,33
   33 X(I,3)=0.0
    8 CONTINUE
C CALCULATION OF X TRANSPOSE X INVERSE
      DO 11 I=1,3
      DO 11 J=1,3
      XTX(I,J)=0.0
   11 XTXIN(I,J)=0.0
      FNTOT=NTOT
      XTX(1,1)=FNTOT
      DO 1001 I=1,NTOT
      XTX(1,2)=XTX(1,2)+X(I,1)*X(I,2)
      XTX(1,3)=XTX(1,3)+X(I,1)*X(I,3)
      XTX(2,2)=XTX(2,2)+X(I,2)**2
      XTX(3,3)=XTX(3,3)+X(I,3)**2
 1001 XTX(2,3)=XTX(2,3)+X(I,2)*X(I,3)
      DO 1002 I=1,2
      DO 1002 J=1,3
 1002 XTX(J,I)=XTX(I,J)
      DET=XTX(1,1)*(XTX(2,2)*XTX(3,3)-XTX(2,3)**2)
      DET=DET-XTX(2,1)*(XTX(1,2)*XTX(3,3)-XTX(3,2)*XTX(1,3))
      DET=DET+XTX(3,1)*(XTX(1,2)*XTX(2,3)-XTX(2,2)*XTX(1,3))
      XTXIN(1,1)=(XTX(2,2)*XTX(3,3)-XTX(2,3)**2)/DET
      XTXIN(2,2)=(XTX(1,1)*XTX(3,3)-XTX(1,3)**2)/DET
      XTXIN(3,3)=(XTX(1,1)*XTX(2,2)-XTX(1,2)**2)/DET
      XTXIN(1,2)=-(XTX(2,1)*XTX(3,3)-XTX(3,1)*XTX(2,3))/DET
      XTXIN(1,3)=(XTX(2,1)*XTX(3,2)-XTX(3,1)*XTX(2,2))/DET
      XTXIN(2,3)=-(XTX(1,1)*XTX(3,2)-XTX(3,1)*XTX(1,2))/DET
      DO 1003 I=1,2
      DO 1003 J=1,3
 1003 XTXIN(J,I)=XTXIN(I,J)
      DO 13 I=1,3
      THETA(I)=0.0
   13 XTY(I)=0.0
      DO 14 J=1,3
      DO 14 I=1,NTOT
   14 XTY(J)=XTY(J)+X(I,J)*Y(I)
C CALCULATION OF THETA... L HAT AND DELTA  HAT
      DO 15 J=1,3
      DO 15 I=1,3
   15 THETA(J)=THETA(J)+XTXIN(J,I)*XTY(I)
      YTY=0.0
C CALCULATION OF THE RESIDUAL VARIANCE
      DO 18 I=1,NTOT
   18 YTY=YTY+Y(I)**2
      DO 31 I=1,3
   31 XTY(I)=0.0
      DO 19 J=1,3
      DO 19 I=1,3
   19 XTY(J)=XTY(J)+THETA(I)*XTX(I,J)
      FITVAR=0.0
      DO 20 I=1,3
   20 FITVAR=FITVAR+XTY(I)*THETA(I)
      S=YTY-FITVAR
      S=S/(FNTOT-3.0)
C CALCULATION OF THE STANDARD ERRORS OF DELTA AND MU
      SMU=S*XTXIN(1,1)
      SDELTA=S*XTXIN(3,3)
      SMU=SQRT(SMU)
      SDELTA=SQRT(SDELTA)
      TMU=THETA(1)/SMU
      TDELTA=THETA(3)/SDELTA
C CALCULATION OF THE POSTERIOR DISTRIBUTION... LOGS ARE USED TO PREVENT
C OVERFLOW
      SK=ALOG(S)
      DET=ALOG(DET)
      H=(-.5*DET)-(.5*(FNTOT-3.0)*SK)
      H=.4342945*H
      JK=JK+1
      XOUT(JK,1)=G
      XOUT(JK,2)=S
      XOUT(JK,3)=THETA(2)
      XOUT(JK,4)=THETA(1)
      XOUT(JK,5)=SMU
      XOUT(JK,6)=TMU
      XOUT(JK,7)=THETA(3)
      XOUT(JK,8)=SDELTA
      XOUT(JK,9)=TDELTA
      IF(NTOT-30)1004,1005,1005
 1005 DELCON(JK,1)=THETA(3)-2.58*SDELTA
      DELCON(JK,2)=THETA(3)-1.96*SDELTA
      DELCON(JK,3)=THETA(3)
      DELCON(JK,4)=THETA(3)+1.96*SDELTA
      DELCON(JK,5)=THETA(3)+2.58*SDELTA
 1004 NIT=NIT+1
      PD(NIT)=H
C INCREMENT GAMMA BY .01 AND ITERATE
      G=G+1.00000000E-02
      IF(NIT-199)30,30,26
   30 GO TO 25
   26 CONTINUE
C FIND MAXIMUM VALUE OF THE POSTERIOR
      FIN=PD(1)
      DO 506 I=2,199
      IF(FIN-PD(I))505,506,506
  505 FIN=PD(I)
  506 CONTINUE
C RESCALE POSTERIOR BY DIVIDING ALL VALUES OF THE POSTERIOR BY THE MAX VALUE
      DO 507 I=1,199
      PD(I)=PD(I)-FIN
      YY=ABS(PD(I))
      IF(YY-35.0)509,508,508
  508 PD(I)=0.0
      GO TO 507
  509 PD(I)=PD(I)/.4342945
      PD(I)=EXP(PD(I))
  507 CONTINUE
C CONVERT AREA OF THE POSTERIOR DISTRIBUTION TO UNIT AREA BY METHOD OF TRAPEZO
      AREA=0.0
      DO 511 I=2,199
      II=I-1
  511 AREA=AREA+.005*(PD(I)+PD(II))
      DO 512 I=1,199
  512 PD(I)=PD(I)/AREA
      DO 513 I=1,199
  513 XOUT(I,10)=PD(I)
      DO 514 I=1,199
  514 WRITE(6,23)(XOUT(I,J),J=1,10)
   23 FORMAT(3X,1F5.2,4X,9E11.3)
      DO 700 I=1,199
      XX(I,1)=XOUT(I,9)
  700 XX(I,2)=XOUT(I,10)
      M=199
      N=2
      ISCALE=1
C TITLE FOR  PLOT OF THE POSTERIOR DISTRIBUTION AND FOR STUDENT T
      DO 760 I=1,5
      FNAM(1,I)=G1(I)
  760 FNAM(2,I)=G2(I)
      CALL PLOT(M,N,ISCALE,XX,FNAM)
      N=5
C PLOT CONFIDENCE INTERVALS AROUND DELTA
      IF(NTOT-30)1007,1006,1006
 1006 DO 1008 I=1,200
      DO 1008 J=1,5
 1008 XX(I,J)=DELCON(I,J)
      DO 756 I=1,5
      FNAM(1,I)=G3(I)
      FNAM(2,I)=G4(I)
      FNAM(3,I)=G5(I)
      FNAM(4,I)=G6(I)
  756 FNAM(5,I)=G7(I)
      WRITE(6,1016)
 1016 FORMAT(//34H CONFIDENCE INTERVALS AROUND DELTA)
      WRITE(6,1017)
 1017 FORMAT(/3X,5HGAMMA,2X,52H  LOWER 99   LOWER 95   DELTA    UPPER 95
     1   UPPER 99)
      DO 1019 I=1,199
 1019 WRITE(6,1018)XOUT(I,1),(DELCON(I,J),J=1,5)
 1018 FORMAT(3X,1F5.2,1X5E11.3)
```

```
      WRITE(6,1015)
 1015 FORMAT(//1X,47H GRAPH OF CONFIDENCE INTERVALS AROUND DELTA HAT)
      CALL PLOT(M,N,ISCALE,XX,FNAM)
 1007 NB=NB-1
      IF(NB)38,38,39
   38 CONTINUE
      CALL EXIT
      END
      SUBROUTINE CORREL (N1,N2,Z)
      DIMENSION Z(500),RLAG(400),Y(500),XX(200,10),FNAM(5,5)
      DIMENSION G8(5),G9(5)
      DATA G8/20HPRE TREATMENT DATA  /
      DATA G9/20HPOST TREATMENT DATA /
      NTOT=N1+N2
C PREPARATION OF TITLE FOR CORRELOGRAM PLOT
      DO 804 JJJ=1,4
      IND=JJJ
      IF(IND-2)805,806,807
  807 IF(IND-4)805,806,806
  805 CONTINUE
      DO 757 I=1,5
  757 FNAM(1,I)=G8(I)
      GO TO 808
  806 CONTINUE
      DO 758 I=1,5
  758 FNAM(1,I)=G9(I)
  808 CONTINUE
      GO TO (809,810,811,812),IND
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR PRE TREATMENT DATA
  809 NLAG=N1
      NLOW=1
      NTOP=N1
      WRITE(6,813)
  813 FORMAT(//38H CORRELOGRAM OF PRE TREATMENT RAW DATA)
      GO TO 814
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR POST TREATMENT DATA
  810 NLAG=N2
      NLOW=N1+1
      NTOP=NTOT
      WRITE(6,815)
  815 FORMAT(//39H CORRELOGRAM OF POST TREATMENT RAW DATA)
  814 II=0
      DO 706 I=NLOW,NTOP
      II=II+1
  706 Y(II)=Z(I)
      GO TO 709
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR PRE TREATMENT
C DIFFERENCES  BETWEEN SUCCESSIVE OBSERVATIONS
  811 NLAG=N1-1
      NLOW=1
      NTOP=N1-1
      WRITE(6,800)
  800 FORMAT(//41H CORRELOGRAM OF PRE TREATMENT DIFFERENCES)
      GO TO 816
C PREPARATION OF DATA TO CALCULATE AUTOCORRELATIONS FOR POST TREATMENT
C DIFFERENCES  BETWEEN SUCCESSIVE OBSERVATIONS
  812 NLAG=N2-1
      NLOW=N1+1
      NTOP=NTOT-1
      WRITE(6,803)
  803 FORMAT(//42H CORRELOGRAM OF POST TREATMENT DIFFERENCES)
  816 II=0
      DO 707 I=NLOW,NTOP
      II=II+1
      IK=I+1
  707 Y(II)=Z(IK)-Z(I)
  709 JJ=NLAG*3/4
      DO 716 K=1,JJ
      NUP=NLAG-K
      CROS=0.0
      SUM1=0.0
      SUM2=0.0
      SUMS1=0.0
      SUMS2=0.0
      DO 703 I=1,NUP
      NIND=I+K
      CROS=CROS+Y(I)*Y(NIND)
      SUM1=SUM1+Y(I)
      SUM2=SUM2+Y(NIND)
      SUMS1=SUMS1+Y(I)*Y(I)
 703  SUMS2=SUMS2+Y(NIND)*Y(NIND)
      FNUP=NUP
      DNUM=CROS-(SUM1*SUM2)/FNUP
      DEN=(SUMS1-(SUM1*SUM1)/FNUP)*(SUMS2-(SUM2*SUM2)/FNUP)
      DEN=SQRT(DEN)
      RLAG(K)=DNUM/DEN
      WRITE(6,704)K,RLAG(K)
 704  FORMAT(1X,5H LAG=I3,3X,3H R=F6.3)
 716  CONTINUE
      ISCALE=1
      N=1
      M=JJ
      DO 650 I=1,JJ
  650  XX(I,1)=RLAG(I)
      CALL PLOT(M,N,ISCALE,XX,FNAM)
  804 CONTINUE
      RETURN
      END
      SUBROUTINE PLOT (M,N,ISCALE,XX,FNAM)
C ADAPTED FROM THE SUBROUTINE-GRAPH- OF -PERSUB- WRITTEN BY J.H.WARD,JR.,
C KATHLEEN DAVIS, AND JANICE BUCHHORN,LACKLAND AIR FORCE BASE,TEXAS
C ADAPTED FOR FORTRAN II BY T.O.MAGUIRE,UNIVERSITY OF ILLINOIS
      DIMENSION RAT(10),FND(5,10),ENCRMT(10),PA(120),FMT(3),PB(10)
      DIMENSION XX(200,10),AMAX(10),AMIN(10),B(10),FNAM(5,5),BB(10)
      DATA BB(1),BB(2),BB(3),BB(4),BB(5),BB(6),BB(7),BB(8),BB(9),BB(10)/
     11H1,1H5,1HD,1H5,1H1,1HO,1HX,1HM,1HH,1HA/
      DATA PCROSS,PBLANK,PDASH,PERIOD/1H*,1H ,1H-,1H./
      DATA FMT(1),FMT(2),FMT(3)/1H ,1HO,1H-/
      IF(N-5)200,201,200
 201  DO 276 I=1,5
 276  B(I)=BB(I)
      GO TO 202
 200  DO 277 I=1,3
      II=I+5
 277  B(I)=BB(II)
 202  CONTINUE
  99  FMS=FMT(ISCALE-1)
   98 DO 100 I=1,N
      AMIN(I)=+1.E+37
 100  AMAX(I)=-1.E+37
```

```
C SEARCH FOR MAXIMA AND MINIMA
      ABMIN=AMIN(1)
      ABMAX=AMAX(1)
      DO 101 I=1,M
      DO 101 J=1,N
      COMP=XX(I,J)-AMIN(J)
      IF(COMP)4002,4003,4003
 4002 AMIN(J)=XX(I,J)
 4003 COMP=XX(I,J)-AMAX(J)
      IF(COMP)101,101,4004
 4004 AMAX(J)=XX(I,J)
  101 CONTINUE
      IF(N-5)250,251,250
  251 DO 252 J=1,5
      AMAX(J)=AMAX(5)
  252 AMIN(J)=AMIN(1)
  250 CONTINUE
      N1=N
      DO 108 J=1,N
C COMPUTE RESOLUTION OF GRAPH
  108 RAT(J)=(AMAX(J)-AMIN(J))/110.
      DO 110   J=1,N
      ENCRMT(J)=(AMAX(J)-AMIN(J))/4.
      FND(1,J)=AMIN(J)+.05
      COMP=AMIN(J)
      IF(COMP)4005,4006,4006
 4005 FND(1,J)=FND(1,J)-.10
 4006 FND(5,J)=AMAX(J)-.05
      COMP=AMAX(J)
      IF(COMP)4007,4008,4008
C PREPARE ORDINATE LABELS
 4007 FND(5,J)=FND(5,J)-.10
 4008 FND(2,J)=AMIN(J)+ENCRMT(J)+.05
      COMP=FND(2,J)
      IF(COMP)4009,4010,4010
 4009 FND(2,J)=FND(2,J)-.10
 4010 FND(3,J)=AMIN(J)+(ENCRMT(J)*2.)+.05
      COMP=FND(3,J)
      IF(COMP)4011,4012,4012
 4011 FND(3,J)=FND(3,J)-.10
 4012 FND(4,J)=AMAX(J)-ENCRMT(J)+.05
      COMP=FND(4,J)
      IF(COMP)4013,110,110
 4013 FND(4,J)=FND(4,J)-.10
  110 CONTINUE
C PRINT LEFT HAND LABELS
      WRITE(6,7)((FND(I,J),I=1,5),B(J),J=1,N1)
    7 FORMAT(///1XF6.1,21XF6.1,21XF6.1,21XF6.1,21XF6.1,3XA1)
C PLOT LEFT HAND MARGIN
      DO 4014 I=1,59
 4014 PA(I)=PDASH
      WRITE(6,8)(PA(I),I=1,59)
    8 FORMAT(1X,59A2)
      DO 140 I=1,M
      DO 4015 II=1,120
 4015 PA(II)=PBLANK
      DO 121 IX=27,81,27
  121 PA(IX)=PERIOD
      NCOMP=I
 4016 NCOMP=NCOMP-10
```

```
      IF(NCOMP)124,4017,4016
 4017 CONTINUE
      DO 123 IX=6,120,2
 123  PA(IX)=PDASH
C RESCALE DATA POINTS
 124  DO 135 K=1,N
      ZL=(XX(I,K)-AMIN(K))/RAT(K)+1.0
      L=ZL
      IF(L-1)6018,6019,6019
 6018 L=1
 6019 IF(110-L)6020,6021,6021
 6020 L=110
 6021 IF(PA(L)-PBLANK)6022,130,6022
 6022 IF(PA(L)-PERIOD)6024,130,6024
 6024 IF(PA(L)-PDASH)6025,130,6025
 6025 PA(L)=PCROSS
      GO TO 135
  130 PB(K)=B(K)
      PA(L)=PB(K)
  135 CONTINUE
      IF(ISCALE-1)6023,136,6023
    6 FORMAT(A1)
 6023 WRITE(6,6)FMS
C PLOT DATA POINTS
  136 WRITE(6,2)I,PERIOD,(PA(J),J=1,110),PERIOD,I
    2 FORMAT(1X,I3,A1,110A1,A1,I3)
  140 CONTINUE
      DO 6026 I=1,59
 6026 PA(I)=PDASH
      WRITE(6,8)(PA(I),I=1,59)
      WRITE(6,7)((FND(I,J),I=1,5),B(J),J=1,N1)
  142 WRITE(6,3)
    3 FORMAT(1H0,10X,16HPLOT DESCRIPTION/1H0,7X,5HTITLE,10X,9HCHARACTER,
     14X,7HMINIMUM,4X,7HMAXIMUM,4X,10HRESOLUTION)
      DO 122 J=1,N
 122  WRITE(6,4)(FNAM(J,K),K=1,5),B(J),AMIN(J),AMAX(J),RAT(J)
   4  FORMAT(1X,5A4,2X,A1,7XF8.3,3XF8.3,4XF8.3)
      RETURN
      END
```